

187514

STIC-Biotech/ChemLib

From: Swope, Sheridan
Sent: Thursday, April 27, 2006 4:33 PM
To: STIC-Biotech/ChemLib
Subject: 10/825,632

For 10/825,632, pls search and interference search:

SID 2, oligo search, against the NT and AA databases.

SID 4, regular search, against the NT and AA databases.

SID 6, regular search, against the NT and AA databases.

SID 8, regular search, against the NT and AA databases.

Not a tag of SID 2
Sheridan Swope, Ph.D.
Primary Patent Examiner
AU 1656/Recombinant Enzymes
571-272-0943 (voice)
E02B71 Remsen Bld (Office)
E03C70 Remsen Bld (Mailbox)

CRFE

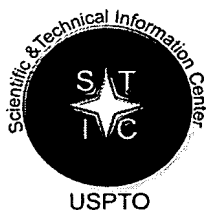
BEST AVAILABLE COPY

Searcher: _____
Searcher Phone: _____
Date Searcher Picked up: _____
Date completed: _____
Searcher Prep Time: _____
Online Time: _____

Type of Search
NA# _____ AA# _____
S/L: _____ Oligomer: _____
Encode/Transl: _____
Structure #: _____ Text: _____
Inventor: _____ Litigation: _____

Vendors and cost where applicable
STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: _____
WWW/Internet: _____
Other (Specify): _____

THIS PAGE BLANK (USPTO)



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 187514

TO: Sheridan Swope
Location: REM-2B71&3C70
Art Unit: 1656
Tuesday, May 09, 2006
Case Serial Number: 10/825632

From: Kristine Hensle
Location: Biotech-Chem Library
REM-1B69
Phone: (571) 272-4161

Kristine. Hensle@uspto.gov

Search Notes

Examiner Swope,

See attached results.

If you have any questions about this search feel free to contact me at any time.

Thank you for using STIC search services!

Kristine Hensle
Librarian
STIC Biotech/Chem Library
(571)272-4161



THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 4, 2006, 04:50:59 ; Search time 15358 Seconds
(without alignments)
11547.830 Million cell updates/sec

Title: US-10-825-632-2
Perfect score: 3120
Sequence: 1 aagtgttaagctccgagg.....agaattactaaaaaaaaa 3120

Scoring table: oligo Nuc
Gapop 60.0, Gapext 60.0

Searched: 5883141 seqs, 28421725653 residues

Word size: 1

Total number of hits satisfying chosen parameters: 11765110

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba:*
2: gb_in:*
3: gb_env:*
4: gb_on:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pr:*
9: gb_ro:*
10: gb_sts:*
11: gb_sy:*
12: gb_un:*
13: gb_vi:*
14: gb_htg:*
15: gb_pl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3120	100.0	3120	6	AR651452 Sequence
2	3120	100.0	3127	8	AF221634 Homo sapi
3	3003	96.2	3106	6	AX342633 Sequence
4	2957	94.8	3143	6	AX354793 Sequence
5	2889	92.6	4829	6	AR631283 Sequence
6	2889	92.6	4829	6	AX608735 Sequence
7	2767	88.7	4535	8	BC040203 Homo sapi
8	2756	88.3	3030	8	AY354202 Homo sapi
9	2547	81.6	2649	6	AX354795 Sequence
10	2495	80.0	2656	8	AY172659 Homo sapi
11	2495	80.0	2671	6	AR631278 Sequence
12	2495	80.0	2671	6	AX608725 Sequence
13	2315	74.2	2797	6	AR448400 Sequence
14	2128	68.2	4523	6	AR631281 Sequence
15	2128	68.2	4523	6	AX608731 Sequence
16	2128	68.2	4676	6	AR631287 Sequence
17	2128	68.2	4676	6	AX608743 Sequence
18	2116	67.8	2842	6	AX405770 Sequence

19	2081	66.7	3125	8	BC030688	Homo sapi
20	1984	63.6	4685	6	AR631288	Sequence
21	1984	63.6	4685	6	AX608745	Sequence
22	1963	62.9	2778	8	AK000290	Homo sapi
23	1450	46.5	2161	6	BD157001	Primer fo
24	1450	46.5	2161	6	AX878058	Sequence
25	1450	46.5	2161	8	AK027826	Sequence
26	1377	44.1	2668	6	AX405771	Homo sapi
27	1191	38.2	4309	6	AR631284	Sequence
28	1191	38.2	4309	6	AX608737	Sequence
29	1164	37.3	2510	6	AX338497	Sequence
30	1023	32.8	1669	6	AR651454	Sequence
31	1023	32.8	1669	8	AF221636	Sequence
32	878	28.1	2098	6	CQ721272	Homo sapi
33	877	28.1	1356	6	AR631282	Sequence
34	877	28.1	1356	6	AX608733	Sequence
35	790	25.3	1197	6	AR651453	Sequence
36	790	25.3	1197	8	AF221635	Homo sapi
37	789	25.3	1083	6	AR651455	Sequence
38	789	25.3	1083	8	AF221637	Homo sapi
39	760	24.4	832	6	AR631286	Sequence
40	760	24.4	832	6	AX608741	Sequence
41	739	23.7	873	6	AX083130	Sequence
42	708	22.7	735	6	AX524942	Sequence
43	700	22.4	753	6	AX524935	Sequence
44	672	21.5	1002	6	AX524939	Sequence
45	501	16.1	631	6	BD149852	Primer fo

ALIGNMENTS

RESULT 1	AR651452	AR651452	3120 bp	DNA	linear	PAT 20-APR-2005
LOCUS	Sequence 2 from patent US 6881564.					
DEFINITION	Sequence 2 from patent US 6881564.					
ACCESSION	AR651452					
VERSION	AR651452.1	GI:62795938				
KEYWORDS	Unknown.					
SOURCE	Unknown.					
ORGANISM	Unclassified.					
REFERENCE	1 (bases 1 to 3120)					
AUTHORS	Abbott, C.A. and Gorrell, M.D.					
TITLE	Dipeptidyl peptidases					
JOURNAL	Patent: US 6881564-A 2 19-APR-2005;					
	The University of Sydney; Sydney;					
AUX;						
FEATURES	Location/Qualifiers					
source	1..3120					
	/organism="unknown"					
	/mol_type="genomic DNA"					

ORIGIN	Query Match	100.0%;	Score 3120;	DB 6;	Length 3120;
	Best Local Similarity	100.0%;	Pred. No. 0;		
	Matches 3120;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	AAGTGTAAAGCTTCGAGGCAAGCCGCTGCTACTGCGCGCGCTGCTTCTTACTGCGG	60		
Db	1	AAGTGTAAAGCTTCGAGGCAAGCCGCTGCTACTGCGCGCGCTGCTTCTTACTGCGG	60		
Qy	61	CGTTTCGCGCGCTGGTTGTCTACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	120		
Db	61	CGTTTCGCGCGCTGGTTGTCTACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	120		
Qy	121	GAGTGGAGGGCGCGCAGCATGAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	180		
Db	121	GAGTGGAGGGCGCGCAGCATGAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	180		
Qy	181	TCCGGCG	240		
Db	181	TCCGGCG	240		

QY	241	CTGGTGTGAGATATTTGAACTGCGGACTGTGAGGAGATATTTGATCATCAGGATCGG	300
DB	241	CTGGTGTGAGATATTTGAACTGCGGACTGTGAGGAGATATTTGATCATCAGGATCGG	300
QY	301	CCTAAATGGAGCCTTTTATGTGTGAGCGGTATCTGGAGTCAGCTTAAAGACTGCTT	360
DB	301	CCTAAATGGAGCCTTTTATGTGTGAGCGGTATCTGGAGTCAGCTTAAAGACTGCTT	360
QY	361	GCCGATACCGAGAAATATCATGGCTACATGATGGCTAAGGCACCACTGATTTTCAATGTTT	420
DB	361	GCCGATACCGAGAAATATCATGGCTACATGATGGCTAAGGCACCACTGATTTTCAATGTTT	420
QY	421	GTGAAGAGGAATGATCCAGATGGACCTCAATTCAGACAGAACTTATTACCTTGGCATGTCT	480
DB	421	GTGAAGAGGAATGATCCAGATGGACCTCAATTCAGACAGAACTTATTATTACCTTGGCATGTCT	480
QY	481	GGTGAGAACAGAGAAATATACACTGTTTATTCTGAAATTTCCCAAACTATCAATAGAGCA	540
DB	481	GGTGAGAACAGAGAAATATACACTGTTTATTCTGAAATTTCCCAAACTATCAATAGAGCA	540
QY	541	GCAGTCTTAATGCTCTCTTGGAGCCTCTTTTGGATCTTTTTCAGGCAACACTGGACTAT	600
DB	541	GCAGTCTTAATGCTCTCTTGGAGCCTCTTTTGGATCTTTTTCAGGCAACACTGGACTAT	600
QY	601	GGAATGTATTCTCGAGAAGAACTATTAAGAGAGAAAGAAACGATTTGGAAACAGTCGGA	660
DB	601	GGAATGTATTCTCGAGAAGAACTATTAAGAGAGAAAGAAACGATTTGGAAACAGTCGGA	660
QY	661	ATTGCTTCTTACGATATCACCAAGGAAGTGGAACATTTCTGTTTCAAGCCGGTAGTGA	720
DB	661	ATTGCTTCTTACGATATCACCAAGGAAGTGGAACATTTCTGTTTCAAGCCGGTAGTGA	720
QY	721	ATTTATCACCTTAAGATGGAGGCGCAAGGATTTAGCAACACCTTTAAGGCCCAAT	780
DB	721	ATTTATCACCTTAAGATGGAGGCGCAAGGATTTAGCAACACCTTTAAGGCCCAAT	780
QY	781	CTAGTGAAACTAGTTGTCTCCAAACATACGGATGGATCCAAATTTATGCCCCGCTGATCCA	840
DB	781	CTAGTGAAACTAGTTGTCTCCAAACATACGGATGGATCCAAATTTATGCCCCGCTGATCCA	840
QY	841	GACTGGATTCGTTTTATACATAGCAACGATATTTGGATATCTAAACATCGTAACAGAGAA	900
DB	841	GACTGGATTCGTTTTATACATAGCAACGATATTTGGATATCTAAACATCGTAACAGAGAA	900
QY	901	GAAAGGAGACTCACTTATGTGCACAATGAGCTAGCCAAACATGGAAGAGATGCCAGATCA	960
DB	901	GAAAGGAGACTCACTTATGTGCACAATGAGCTAGCCAAACATGGAAGAGATGCCAGATCA	960
QY	961	GCTGGAGTCGCTACCTTTGTTCTTCCAAAGAGAAATTTGATAGATATTTCTGGCTATTGGTGG	1020
DB	961	GCTGGAGTCGCTACCTTTGTTCTTCCAAAGAGAAATTTGATAGATATTTCTGGCTATTGGTGG	1020
QY	1021	TGTCCTAAAGCTGAAACAACTCCAGTGGTGTGTAATAATTTCTAGAAATTTCTATATGAAGAA	1080
DB	1021	TGTCCTAAAGCTGAAACAACTCCAGTGGTGTGTAATAATTTCTAGAAATTTCTATATGAAGAA	1080
QY	1081	AATGATGAATCTCAGGTGGAAATTTATCATGTTTACATCCCTATGTTGGAAACAAGGAGG	1140
DB	1081	AATGATGAATCTCAGGTGGAAATTTATCATGTTTACATCCCTATGTTGGAAACAAGGAGG	1140
QY	1141	CGAGATTCATCCGTTATCTTAAACAGGTACAGCAATCTTAAAGTCACTTTTAAAGATG	1200
DB	1141	CGAGATTCATCCGTTATCTTAAACAGGTACAGCAATCTTAAAGTCACTTTTAAAGATG	1200
QY	1201	TCAGAAATATGATGCTGAGGAAGGATCATAGATGTCATAGATAAGGAACATAAT	1260
DB	1201	TCAGAAATATGATGCTGAGGAAGGATCATAGATGTCATAGATAAGGAACATAAT	1260
QY	1261	CAACCTTTTGAGATCTTATTTGAAGAGGTTGAATATATTGGCAGAGCTGGATGGACTCCT	1320
DB	1261	CAACCTTTTGAGATCTTATTTGAAGAGGTTGAATATATTGGCAGAGCTGGATGGACTCCT	1320

QY	1321	GAGGAAATATATCTTGGTCCATCTTACTAGATCCGCTCCAGACTCCGCTACAGATAGTG	1380
DB	1321	GAGGAAATATATCTTGGTCCATCTTACTAGATCCGCTCCAGACTCCGCTACAGATAGTG	1380
QY	1381	TTGATCTCACCTGAAATTTATTTATCCAGTAGAAGATGATTTATGGAAGGAGAGACTC	1440
DB	1381	TTGATCTCACCTGAAATTTATTTATCCAGTAGAAGATGATTTATGGAAGGAGAGACTC	1440
QY	1441	ATTGAGTCAGTCCCTGATTTCTGFGAGCCCACTAAATTTATCTATGAAAGAAACAACAGACATC	1500
DB	1441	ATTGAGTCAGTCCCTGATTTCTGFGAGCCCACTAAATTTATCTATGAAAGAAACAACAGACATC	1500
QY	1501	TGGATAAATATCCATGACATCTTTTCATGTTTTTCCCAAGGTCACGAAAGAGAAATTCAG	1560
DB	1501	TGGATAAATATCCATGACATCTTTTCATGTTTTTCCCAAGGTCACGAAAGAGAAATTCAG	1560
QY	1561	TTTATTTTCCCTCTGAAATGCAAAACAGGTTTCCGCTCATTTATACAAATTTACATCTATT	1620
DB	1561	TTTATTTTCCCTCTGAAATGCAAAACAGGTTTCCGCTCATTTATACAAATTTACATCTATT	1620
QY	1621	TTAAAGGAAAGCAAAATATAAAGCATCCAGTGGTGGCTGCCTGCCCAAGTATTCAAG	1680
DB	1621	TTAAAGGAAAGCAAAATATAAAGCATCCAGTGGTGGCTGCCTGCCCAAGTATTCAAG	1680
QY	1681	TGCTCTATCAAGAGGAGATAGCAATTTACAGTGTGTAATGGAAGTCTTTGGCGGCAT	1740
DB	1681	TGCTCTATCAAGAGGAGATAGCAATTTACAGTGTGTAATGGAAGTCTTTGGCGGCAT	1740
QY	1741	GGATCTAATATCCAACTTGATGTAAGTCAGAGGCTGTATATTTTGAAGGCCACCAAGAC	1800
DB	1741	GGATCTAATATCCAACTTGATGTAAGTCAGAGGCTGTATATTTTGAAGGCCACCAAGAC	1800
QY	1801	TCCCTTTTAGAGATCACCTGTACGTAGTCAGTACGTAAGTAAATCTTGGAGAGGTGACAAAG	1860
DB	1801	TCCCTTTTAGAGATCACCTGTACGTAGTCAGTACGTAAGTAAATCTTGGAGAGGTGACAAAG	1860
QY	1861	CTGACTGACCGTGGCTACTCAATTTCTGCTGCATCAGTCAGCACTGFGACTTCTTTATA	1920
DB	1861	CTGACTGACCGTGGCTACTCAATTTCTGCTGCATCAGTCAGCACTGFGACTTCTTTATA	1920
QY	1921	AGTAAGTATATAGTAACAGAGAAATCCACACTGTGTCTTCTTCAAGCTATCAAGTCTCT	1980
DB	1921	AGTAAGTATATAGTAACAGAGAAATCCACACTGTGTCTTCTTCAAGCTATCAAGTCTCT	1980
QY	1981	GAGATGACCCAACTTTGCAAAAACAAAGAAATTTTGGGCGCACCATTTTGGATTCAGCAGGT	2040
DB	1981	GAGATGACCCAACTTTGCAAAAACAAAGAAATTTTGGGCGCACCATTTTGGATTCAGCAGGT	2040
QY	2041	CCTCTTCTGACTATCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGAATTTACA	2100
DB	2041	CCTCTTCTGACTATCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGAATTTACA	2100
QY	2101	TTGTATGGGATGCTCTACAAGCTCATGATCTACAGCTGGAAGAAATATCTTACTGTG	2160
DB	2101	TTGTATGGGATGCTCTACAAGCTCATGATCTACAGCTGGAAGAAATATCTTACTGTG	2160
QY	2161	CTGTTTCAATATCGTGGTCTCAGGTGCAATGTTGTTGTAATATCGGTTTAAAGGAGTCAAG	2220
DB	2161	CTGTTTCAATATCGTGGTCTCAGGTGCAATGTTGTTGTAATATCGGTTTAAAGGAGTCAAG	2220
QY	2221	TATTTCCGCTTGAATACCTAGCTCTCTAGGTATGTTGTTGTTAGTATAGACACAGG	2280
DB	2221	TATTTCCGCTTGAATACCTAGCTCTCTAGGTATGTTGTTGTTAGTATAGACACAGG	2280
QY	2281	GGATCTGTGCAACGAGGCTTAAATTTGAAAGGCGCTTTAAATATAAATGGGTCAAATA	2340
DB	2281	GGATCTGTGCAACGAGGCTTAAATTTGAAAGGCGCTTTAAATATAAATGGGTCAAATA	2340
QY	2341	GAAATTTGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTCAATGAC	2400
DB	2341	GAAATTTGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTCAATGAC	2400
QY	2401	TTAGATCGTGGGCACTCCAGCGCTCTATGAGGATACCTCTCCCTGATGGCATTA	2460

FEATURES		Location/Qualifiers	
source		1..3127	
		/organism="Homo sapiens"	
		/mol_type="mRNA"	
		/db_xref="taxon:9606"	
		/chromosome="15"	
		/map="15q22"	
gene		1..3127	
		/gene="DPP8"	
CDS		214..2862	
		/gene="DPP8"	
		/note="serine protease; member of prolyl oligopeptidase S9b family belonging to enzyme clan SC; alternatively spliced forms of this mRNA are found in AF221635, AF221636 and AF221637"	
		/codon_start=1	
		/product="dipeptidyl peptidase 8"	
		/protein_id="AAG29766.1"	
		/db_xref="GI:11095188"	
		/translation="MAAAETBQLGVEIFETADCBENIESQDRPKLEPFYVERYSWSQ LKLLADTRKYHGMYMAKAPHDFMFVKRNDPDGPHSDRIYYLAMSGENRNTLPYSEI PKTINRAAVLMLSWKPLLDLFOATLDYGMYSREELLRRKRIGTVGLASDYHQSGG TELFQAGSGIYHVKGDGPGQFTQOPLRPNLVETSCPNRMOPKLPADPDWIAFTHSN DIWISNITREERRLTYVHNELANMEEDARSAGVATFVLOEEDRYSGYWCPCKAETT PSQKILIRLDYEENDESEVEIHHVTSPLMETRRADSFRYPKGTANPKVTFKMSIEMI DAEGRIIDVDKELIQPFILFEGVEYIARAGMTPEGYAMSILLDRSOTRLQIVLIS PELFIPEVDDVDMERQRLIESVPSVTLPIIYEETDIIWINIHDIHFVPPSGHEEIEF IFASECTGPRHLYKITSILKESKYKRSSGGLPAPSDPKCPIKEBIAITSGEWEVLGR HGSNIQVDEVRRLVYFEGTKDSPLEHLLYVSYVNPGEVTRLTDRGYSHSCISQHCND FFTSKYSNQKNPHCVSLYKLSPEDDPTCTCKEFWATILDSAGPLDTPPEIIFSFES VVVIDNRGSCHRGLKPEGAFKYKMGQIEIDDOVEGLQYLASRYDFIDLDRGVHGWSY GGYSLSLMALMORSIDIIRVALAGAPVTLWIFYDTGYTERVMGHPDQNEOGYVLSVAMQ AEPFSPFNRLLLHGFLENVHFAHTSILLSFLVRACKPYDLQIYPERHSIRVPES GEHYELHLHLHQENLGSRIAALKVI"	
ORIGIN			
Query Match		100.0%; Score 3120; DB 8; Length 3127;	
Best Local Similarity		100.0%; Pred. No. 0;	
Matches 3120; Conservative		0; Mismatches 0; Indels 0; Gaps 0;	
QY	1	AAGTGTCTAAAGCCTCCGAGGCCAAGCGCGCTGCTACTGCGCGCGCTGCTTCTTAGTGCGG	60
DB	1	AAGTGTCTAAAGCCTCCGAGGCCAAGCGCGCTGCTACTGCGCGCGCTGCTTCTTAGTGCGG	60
QY	61	CGTTTCGCCGCTGGGTTGTTCACCGCGCGCGCGCGCGAGGAAGCCACTGCAACACGAGACCG	120
DB	61	CGTTTCGCCGCTGGGTTGTTCACCGCGCGCGCGCGCGAGGAAGCCACTGCAACACGAGACCG	120
QY	121	GAGTGGAGGCGCGCGCGAGCATGAAGCGCGCGCGCGCTCCATAGCGCACTGTCGGGACGG	180
DB	121	GAGTGGAGGCGCGCGCGAGCATGAAGCGCGCGCGCGCTCCATAGCGCACTGTCGGGACGG	180
QY	181	TCCGGGCGGGCGCGGGGGAGGAAAATGCAACATGGCAGCAGCAATCGAAACACAGACAG	240
DB	181	TCCGGGCGGGCGCGGGGGAGGAAAATGCAACATGGCAGCAGCAATCGAAACACAGACAG	240
QY	241	CTGGGTGTTCAGATATTTGAAACTGCGGACTGTGAGGAGAAATATTGAATCAAGATCGG	300
DB	241	CTGGGTGTTCAGATATTTGAAACTGCGGACTGTGAGGAGAAATATTGAATCAAGATCGG	300
QY	301	CCTAAATTGGAGCCCTTTTATGTTGAGCGGCTATTCCTGGAGTTCAGCTTAAAAAGCTGCTT	360
DB	301	CCTAAATTGGAGCCCTTTTATGTTGAGCGGCTATTCCTGGAGTTCAGCTTAAAAAGCTGCTT	360
QY	361	GCCGATACCAAGAAAATATCATGCTACATGATGGCTAAGGCCACCAACATGATTTTATGTTT	420
DB	361	GCCGATACCAAGAAAATATCATGCTACATGATGGCTAAGGCCACCAACATGATTTTATGTTT	420
QY	421	GTGAAGAGGAATGATTCAGATGGACCTCATTTGACAGACAACTTATTACCTTGCATGTCT	480
DB	421	GTGAAGAGGAATGATTCAGATGGACCTCATTTGACAGACAACTTATTACCTTGCATGTCT	480

Qy	481	GGTGAGAA	CAGAGAAATACACTGTTTTTATCTCGAAATTCGAAATTC	CAAAATTC	CAATCAATAGAGCA	540	Db	1561	TTTTATTTTGCCTCTGTAATGCAAAACAGGTTCCGTCATTTATACAAATTTACATCTATT		1561
Db	481	GGTGAGAA	CAGAGAAATACACTGTTTTTATCTCGAAATTCGAAATTC	CAAAATTC	CAATCAATAGAGCA	540	Qy	1621	TTAAAGAAAGCAAAATATAAACGATCCAGTGGTGGGCTGCTCTGCTCCAAGTGATTTCAAG		1620
Qy	541	GCAGTCTTAATGCTCTCTTGGAGCCCTCTTTTGGGATCTTTTTCAGGCAACACCTGGACTAT	600	Db	1621	TTAAAGAAAGCAAAATATAAACGATCCAGTGGTGGGCTGCTCTGCTCCAAGTGATTTCAAG	1680	1621	TTAAAGAAAGCAAAATATAAACGATCCAGTGGTGGGCTGCTCTGCTCCAAGTGATTTCAAG		1680
Db	541	GCAGTCTTAATGCTCTCTTGGAGCCCTCTTTTGGGATCTTTTTCAGGCAACACCTGGACTAT	600	Qy	1681	TGTCCTATCAAGAGAGAGATAGCAATTTACCAGTGGTGAATGGGAAGTTCTTTGGCCGGCAT	1740	1681	TGTCCTATCAAGAGAGAGATAGCAATTTACCAGTGGTGAATGGGAAGTTCTTTGGCCGGCAT		1740
Qy	601	GGAAATGTAATCTCGAGAAGAAGAACTATTTAAGAGAAAGAAACGCAATTTGGAACAGTCGGA	660	Db	1681	TGTCCTATCAAGAGAGAGATAGCAATTTACCAGTGGTGAATGGGAAGTTCTTTGGCCGGCAT	1740	1681	TGTCCTATCAAGAGAGAGATAGCAATTTACCAGTGGTGAATGGGAAGTTCTTTGGCCGGCAT		1740
Db	601	GGAAATGTAATCTCGAGAAGAAGAACTATTTAAGAGAAAGAAACGCAATTTGGAACAGTCGGA	660	Qy	1741	GGATCTAATATCCAAGTGTGAAAGTCAAGAGGCTGGTATATTTTGAAGGCACCAAGAGAC	1800	1741	GGATCTAATATCCAAGTGTGAAAGTCAAGAGGCTGGTATATTTTGAAGGCACCAAGAGAC		1800
Qy	661	ATTGCTTCTTTACGATATACCAAGGAAGTGAACATTTCTGTGTTCAAGCCGGTGTAGTGA	720	Db	1741	GGATCTAATATCCAAGTGTGAAAGTCAAGAGGCTGGTATATTTTGAAGGCACCAAGAGAC	1800	1741	GGATCTAATATCCAAGTGTGAAAGTCAAGAGGCTGGTATATTTTGAAGGCACCAAGAGAC		1800
Db	661	ATTGCTTCTTTACGATATACCAAGGAAGTGAACATTTCTGTGTTCAAGCCGGTGTAGTGA	720	Qy	1801	TCCCTTTAGAGCATCACCTGTACGTAGTACGTAGTACGTAGTAAATCTCGAGAGGTCACAAGG	1860	1801	TCCCTTTAGAGCATCACCTGTACGTAGTACGTAGTAAATCTCGAGAGGTCACAAGG		1860
Qy	721	ATTTATCAGCTAAAGATGGAGGCCACAGGATTTAGCAACACCTTTTAAGGCCCAAT	780	Db	1801	TCCCTTTAGAGCATCACCTGTACGTAGTACGTAGTAAATCTCGAGAGGTCACAAGG	1860	1801	TCCCTTTAGAGCATCACCTGTACGTAGTAAATCTCGAGAGGTCACAAGG		1860
Db	721	ATTTATCAGCTAAAGATGGAGGCCACAGGATTTAGCAACACCTTTTAAGGCCCAAT	780	Qy	1861	CTGACTGACCGTGGCTACTCACAATCTTGTCTGCATCACTGACACTGTGACTTCTTTATA	1920	1861	CTGACTGACCGTGGCTACTCACAATCTTGTCTGCATCACTGACACTGTGACTTCTTTATA		1920
Qy	781	CTAGTGGAAATCTAGTTGTCCTCAACATACGGATGGATCCAAATTTATGCCCCGCTGATCCA	840	Db	1861	CTGACTGACCGTGGCTACTCACAATCTTGTCTGCATCACTGACACTGTGACTTCTTTATA	1920	1861	CTGACTGACCGTGGCTACTCACAATCTTGTCTGCATCACTGACACTGTGACTTCTTTATA		1920
Db	781	CTAGTGGAAATCTAGTTGTCCTCAACATACGGATGGATCCAAATTTATGCCCCGCTGATCCA	840	Qy	1921	AGTAAGTATAGTAAACAGAGAAATCCACACTGTGTGCTCCCTTTACAAGCTATCAAGTCCCT	1980	1921	AGTAAGTATAGTAAACAGAGAAATCCACACTGTGTGCTCCCTTTACAAGCTATCAAGTCCCT		1980
Qy	841	GACTGGATTCGTTTTATACATAGCAACGATATTTTGGATATCTAACATCGTAAACAGAGAA	900	Db	1921	AGTAAGTATAGTAAACAGAGAAATCCACACTGTGTGCTCCCTTTACAAGCTATCAAGTCCCT	1980	1921	AGTAAGTATAGTAAACAGAGAAATCCACACTGTGTGCTCCCTTTACAAGCTATCAAGTCCCT		1980
Db	841	GACTGGATTCGTTTTATACATAGCAACGATATTTTGGATATCTAACATCGTAAACAGAGAA	900	Qy	1981	GAAGATGACCCAACTTTGCAAAAACAAAGGAAATTTTGGGCCACATTTTGGATTCAGAGGT	2040	1981	GAAGATGACCCAACTTTGCAAAAACAAAGGAAATTTTGGGCCACATTTTGGATTCAGAGGT	2040	2040
Qy	901	GAAGAGAGACTCACTTATGTGCACAATGAGCTAGGCCAAACATGGAAGAGATGCCAGATCA	960	Db	1981	GAAGATGACCCAACTTTGCAAAAACAAAGGAAATTTTGGGCCACATTTTGGATTCAGAGGT	2040	1981	GAAGATGACCCAACTTTGCAAAAACAAAGGAAATTTTGGGCCACATTTTGGATTCAGAGGT	2040	2040
Db	901	GAAGAGAGAGACTCACTTATGTGCACAATGAGCTAGGCCAAACATGGAAGAGATGCCAGATCA	960	Qy	2041	CTCTCTCTGACTATATCTCTCCAGAAATTTTCTCTTTTGAAGATGTACTACCTGGAATTTACA	2100	2041	CTCTCTCTGACTATATCTCTCCAGAAATTTTCTCTTTTGAAGATGTACTACCTGGAATTTACA	2100	2100
Qy	961	GCTGGAGTCGCTACCTTTGTTCTTCAAGGAAGAAATTTGATAGATATTCCTGGCTATTTGGTG	1020	Db	2041	CTCTCTCTGACTATATCTCTCCAGAAATTTTCTCTTTTGAAGATGTACTACCTGGAATTTACA	2100	2041	CTCTCTCTGACTATATCTCTCCAGAAATTTTCTCTTTTGAAGATGTACTACCTGGAATTTACA	2100	2100
Db	961	GCTGGAGTCGCTACCTTTGTTCTTCAAGGAAGAAATTTGATAGATATTCCTGGCTATTTGGTG	1020	Qy	2101	TTGTATGGGATGCTCTACAAGCTCATGATCTACAGCCTGGAAGAAATATCTCTACCTGTG	2160	2101	TTGTATGGGATGCTCTACAAGCTCATGATCTACAGCCTGGAAGAAATATCTCTACCTGTG	2160	2160
Qy	1021	TGTCCTCAAAAGCTGAAACCACTCCAGTGGTGTGTAATTTCTAGAAATTTCTATATGAAGAA	1080	Db	2101	TTGTATGGGATGCTCTACAAGCTCATGATCTACAGCCTGGAAGAAATATCTCTACCTGTG	2160	2101	TTGTATGGGATGCTCTACAAGCTCATGATCTACAGCCTGGAAGAAATATCTCTACCTGTG	2160	2160
Db	1021	TGTCCTCAAAAGCTGAAACCACTCCAGTGGTGTGTAATTTCTAGAAATTTCTATATGAAGAA	1080	Qy	2161	CTGTTCATATATGGTGGTCTCTCAGGTGCAATTTGGTGAATATATCGGTTTAAAGAGGATCAAG	2220	2161	CTGTTCATATATGGTGGTCTCTCAGGTGCAATTTGGTGAATATATCGGTTTAAAGAGGATCAAG	2220	2220
Qy	1081	AATCATGATCTGAGGTGGAAATTTATTCATGTTTACATCCCTATGTTGGAAACCAAGGAGG	1140	Db	2161	CTGTTCATATATGGTGGTCTCTCAGGTGCAATTTGGTGAATATATCGGTTTAAAGAGGATCAAG	2220	2161	CTGTTCATATATGGTGGTCTCTCAGGTGCAATTTGGTGAATATATCGGTTTAAAGAGGATCAAG	2220	2220
Db	1081	AATCATGATCTGAGGTGGAAATTTATTCATGTTTACATCCCTATGTTGGAAACCAAGGAGG	1140	Qy	2221	TATTTCCGCTTGAATATACCTTAGCCTCTCTAGGTTATGTGGTTGTAGTGATAGACAAACAGG	2280	2221	TATTTCCGCTTGAATATACCTTAGCCTCTCTAGGTTATGTGGTTGTAGTGATAGACAAACAGG	2280	2280
Qy	1141	GCAGATTCATTCGGTTATCCTTAAACAGGTACAGCAATCTCTAAAGTCACCTTTTAAGATG	1200	Db	2221	TATTTCCGCTTGAATATACCTTAGCCTCTCTAGGTTATGTGGTTGTAGTGATAGACAAACAGG	2280	2221	TATTTCCGCTTGAATATACCTTAGCCTCTCTAGGTTATGTGGTTGTAGTGATAGACAAACAGG	2280	2280
Db	1141	GCAGATTCATTCGGTTATCCTTAAACAGGTACAGCAATCTCTAAAGTCACCTTTTAAGATG	1200	Qy	2281	GGATCTCTCACCGAGGCTTAAATTTTGAAGCGCCTTTAAATATAAAATGGGTCAAATA	2340	2281	GGATCTCTCACCGAGGCTTAAATTTTGAAGCGCCTTTAAATATAAAATGGGTCAAATA	2340	2340
Qy	1201	TCAGAAATATGATGATCTGAAGAGGATCATAGATGTCATAGATAAGGAACTAAT	1260	Db	2281	GGATCTCTCACCGAGGCTTAAATTTTGAAGCGCCTTTAAATATAAAATGGGTCAAATA	2340	2281	GGATCTCTCACCGAGGCTTAAATTTTGAAGCGCCTTTAAATATAAAATGGGTCAAATA	2340	2340
Db	1201	TCAGAAATATGATGATCTGAAGAGGATCATAGATGTCATAGATAAGGAACTAAT	1260	Qy	2341	GAATTTGACCATCAGGTGGAGGACTCCCAATATCTAGCTTCTCGATATGATTTCAATTTGAC	2400	2341	GAATTTGACCATCAGGTGGAGGACTCCCAATATCTAGCTTCTCGATATGATTTCAATTTGAC	2400	2400
Qy	1261	CAACCTTTTGAGATCTTATTTGAAGGATGTAATATTTGCCAGAGCTGGATGACTCCT	1320	Db	2341	GAATTTGACCATCAGGTGGAGGACTCCCAATATCTAGCTTCTCGATATGATTTCAATTTGAC	2400	2341	GAATTTGACCATCAGGTGGAGGACTCCCAATATCTAGCTTCTCGATATGATTTCAATTTGAC	2400	2400
Db	1261	CAACCTTTTGAGATCTTATTTGAAGGATGTAATATTTGCCAGAGCTGGATGACTCCT	1320	Qy	2401	TTAGATCGTGGGCAATCCACGGCTGGTCTATGAGGATACCTCTCCCTGATGGCATT	2460	2401	TTAGATCGTGGGCAATCCACGGCTGGTCTATGAGGATACCTCTCCCTGATGGCATT	2460	2460
Qy	1321	GAGGAAATATGCTGGTCCATCTCTAGATCGGTCAGATCGGCTTACAGATAGTG	1380	Db	2401	TTAGATCGTGGGCAATCCACGGCTGGTCTATGAGGATACCTCTCCCTGATGGCATT	2460	2401	TTAGATCGTGGGCAATCCACGGCTGGTCTATGAGGATACCTCTCCCTGATGGCATT	2460	2460
Db	1321	GAGGAAATATGCTGGTCCATCTCTAGATCGGTCAGATCGGCTTACAGATAGTG	1380	Qy	2461	ATGCAAGGTGCAATATCTTACAGGTTGCTATTTGCTGGGGCCCGACCTCTCTGATGATC	2520	2461	ATGCAAGGTGCAATATCTTACAGGTTGCTATTTGCTGGGGCCCGACCTCTCTGATGATC	2520	2520
Qy	1381	TTGATCTCAGCTGAAATTTATTTCCAGTAGAAGATGATGTTATGGAAGGCAGAGACTC	1440	Db	2461	ATGCAAGGTGCAATATCTTACAGGTTGCTATTTGCTGGGGCCCGACCTCTCTGATGATC	2520	2461	ATGCAAGGTGCAATATCTTACAGGTTGCTATTTGCTGGGGCCCGACCTCTCTGATGATC	2520	2520
Db	1381	TTGATCTCAGCTGAAATTTATTTCCAGTAGAAGATGATGTTATGGAAGGCAGAGACTC	1440	Qy	2521	TTCTATGATACAGGATACAGGAAACGTTATATGGGTCAACCTTGACCAAGATGAACAGGCG	2580	2521	TTCTATGATACAGGATACAGGAAACGTTATATGGGTCAACCTTGACCAAGATGAACAGGCG	2580	2580
Qy	1441	ATTGAGTCAGTCCGCTGATCTCTGAGCCCACTAATTTCTATGAAAGAACCAACAGATC	1500	Db	2521	TTCTATGATACAGGATACAGGAAACGTTATATGGGTCAACCTTGACCAAGATGAACAGGCG	2580	2521	TTCTATGATACAGGATACAGGAAACGTTATATGGGTCAACCTTGACCAAGATGAACAGGCG	2580	2580
Db	1441	ATTGAGTCAGTCCGCTGATCTCTGAGCCCACTAATTTCTATGAAAGAACCAACAGATC	1500	Qy	2581	TATTTACTTAGGATCTGTGGCCATGCAAGCAGAAAAAGTTCCCTCTCTGAAACCAAAATCGTTTA	2640	2581	TATTTACTTAGGATCTGTGGCCATGCAAGCAGAAAAAGTTCCCTCTCTGAAACCAAAATCGTTTA	2640	2640
Qy	1501	TGGATAAATATCCATGACATCTTTCAATGTTTTTCCCAAGTTCAGAAAGAGGAAATTTGAG	1560	Db	2581	TATTTACTTAGGATCTGTGGCCATGCAAGCAGAAAAAGTTCCCTCTCTGAAACCAAAATCGTTTA	2640	2581	TATTTACTTAGGATCTGTGGCCATGCAAGCAGAAAAAGTTCCCTCTCTGAAACCAAAATCGTTTA	2640	2640
Db	1501	TGGATAAATATCCATGACATCTTTCAATGTTTTTCCCAAGTTCAGAAAGAGGAAATTTGAG	1560	Qy	2641	CTGCTCTTACATGATGTTCTCTGGATGAAATGTCCATTTTGGACATACCAAGTATATTAATG	2700	2641	CTGCTCTTACATGATGTTCTCTGGATGAAATGTCCATTTTGGACATACCAAGTATATTAATG	2700	2700
Qy	1561	TTTTATTTTGCCTCTGAAATCAAAACAGGTTTCCGTCATTTATACAAATTTACATCTATT	1620								

Db	2641	CTGCTCTTACATGGTTTCTCTGGATGAGATGTCCATTTTGGACATACACAGTATATTACTG	2700
Qy	2701	AGTTTTTTAGTGGGCTGGAAAGCCATATGATTTACAGATCTATCTCCAGGAGACAC	2760
Db	2701	AGTTTTTTAGTGGGCTGGAAAGCCATATGATTTTACAGATCTATCTCCAGGAGACAC	2760
Qy	2761	AGCATAGAGTTCTCTGAATCGGAGAGAACATTTATGAACCTGCATCTTTTGCACTACCTTCAA	2820
Db	2761	AGCATAGAGTTCTCTGAATCGGAGAGAACATTTATGAACCTGCATCTTTTGCACTACCTTCAA	2820
Qy	2821	GAACACCTTGGATCACGCTATTGCTCTCTAAAGTGATATAAATTTTGACCTGTGTAGAAC	2880
Db	2821	GAACACCTTGGATCACGCTATTGCTCTCTAAAGTGATATAAATTTTGACCTGTGTAGAAC	2880
Qy	2881	TCTCTGGTATACATCTGCTATTAAACCAATGAGGAGTTTAAATCAACAGAGAAACACAGA	2940
Db	2881	TCTCTGGTATACATCTGCTATTAAACCAATGAGGAGTTTAAATCAACAGAGAAACACAGA	2940
Qy	2941	ATTGATCATCATTTTGTATACCTGCCATGTACATCTACTCTGAAATAAATGTGGTG	3000
Db	2941	ATTGATCATCATTTTGTATACCTGCCATGTAAATCATCTACTCTGAAATAAATGTGGTG	3000
Qy	3001	CCATCGAGGGCTCTACGGTTTGGTAGTAATCTTAATACCTTAACCCACATGTCTCAAAA	3060
Db	3001	CCATCGAGGGCTCTACGGTTTGGTAGTAATCTTAATACCTTAACCCACATGTCTCAAAA	3060
Qy	3061	TCAATGATACATATCTCTGAGAGACCCAGCAATACCATAGAAATTTACTTAAAAAAAAA	3120
Db	3061	TCAATGATACATATCTCTGAGAGACCCAGCAATACCATAGAAATTTACTTAAAAAAAAA	3120
RESULT 3			
AX342633			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
JOURNAL			
FEATURES			
source			
ORIGIN			
Query Match			
Best Local Similarity			
Matches 3103; Conservative 0; Mismatches 2; Indels 0; Gaps 0;			
Qy	13	CTCCAGAGCCAGGCGCTGCTACTGCGCGCGTGTCTTCTTAGTGCGCGGTTCGCGGCT	72
Db	2	CTCCAGAGCCAGGCGCTGCTACTGCGCGCGTGTCTTCTTAGTGCGCGGTTCGCGGCT	61
Qy	73	GGGTTGTACCGGCGCGCGCGAGGAGCCACTGCAACAGGACCGAGGTGGAGGGG	132
Db	62	GGGTTGTACCGGCGCGCGCGAGGAGCCACTGCAACAGGACCGAGGTGGAGGGG	121
Qy	133	CGCAGCATGAAGCGCGCGAGCGCGCTCCATAGCGACGTCGCGGACGGTCCGGGCGGGC	192
Db	122	CGCAGCATGAAGCGCGCGAGCGCGCTCCATAGCGACGTCGCGGACGGTCCGGGCGGGC	181

QY	1273	ATTCTATTGAAGAGTTGAATATATATGCGAGAGCTGGATGGACTCTGAGGGGAAATAT	1332
Db	1262	ATTCTATTGAAGAGTTGAATATATATGCGAGAGCTGGATGGACTCTGAGGGGAAATAT	1321
QY	1333	GCTTGGTCCATCTCTACTAGATCGCTCCAGACTCGCCTACAGATAGTGTTCATCTCACCT	1392
Db	1322	GCTTGGTCCATCTCTACTAGATCGCTCCAGACTCGCCTACAGATAGTGTTCATCTCACCT	1381
QY	1393	GAAATATTTATCCCAAGTAGAAGATGATGTTATGGAAGGCGAGAGCTCAATGAGTCAGTG	1452
Db	1382	GAAATATTTATCCCAAGTAGAAGATGATGTTATGGAAGGCGAGAGCTCAATGAGTCAGTG	1441
QY	1453	CCTGATCTGTGAGCCCACTAAATTAATCTATGAAGAAACAAACAGACATCTGGATAAATATC	1512
Db	1442	CCTGATCTGTGAGCCCACTAAATTAATCTATGAAGAAACAAACAGACATCTGGATAAATATC	1501
QY	1513	CATGACATCTTTTCATGTTTTTCCCAAGTTCAGAGAGGAAATTTGAGTTTTATTTTGGC	1572
Db	1502	CATGACATCTTTTCATGTTTTTCCCAAGTTCAGAGAGGAAATTTGAGTTTTATTTTGGC	1561
QY	1573	TCTGAATGCAAAACAGGTTTCGGTCATTTATACAAATTAACATCTATTTTAAAGGAAGC	1632
Db	1562	TCTGAATGCAAAACAGGTTTCGGTCATTTATACAAATTAACATCTATTTTAAAGGAAGC	1621
QY	1633	AAATATAAACATCCAGTGTGGCTGCTGCTCCAGTGATTTTCAAGTGTCCTATCAAA	1692
Db	1622	AAATATAAACATCCAGTGTGGCTGCTGCTCCAGTGATTTTCAAGTGTCCTATCAAA	1681
QY	1693	GAGGAGATAGCAATACCACTGTGTAATGGGAAGTTCTTGGCCGCACTGGAATCAATATC	1752
Db	1682	GAGGAGATAGCAATACCACTGTGTAATGGGAAGTTCTTGGCCGCACTGGAATCAATATC	1741
QY	1753	CAAGTGTATGAAGTCAGAAGGCTGGTATATTTTGAAGGCACCAAGAGCTCCCTTTTAGAG	1812
Db	1742	CAAGTGTATGAAGTCAGAAGGCTGGTATATTTTGAAGGCACCAAGAGCTCCCTTTTAGAG	1801
QY	1813	CATCACCTGTACGTAGTCAGTTACGTAAATCTCTGAGAGGTCGACAAAGGCTGACGACCGT	1872
Db	1802	CATCACCTGTACGTAGTCAGTTACGTAAATCTCTGAGAGGTCGACAAAGGCTGACGACCGT	1861
QY	1873	GGTACTACATCTTCTGCTGCATCAGTCAGCACTGTGACTCTTTTATAGTAAGTATAGT	1932
Db	1862	GGTACTACATCTTCTGCTGCATCAGTCAGCACTGTGACTCTTTTATAGTAAGTATAGT	1921
QY	1933	AACAGAGAATCCACACTGTGTGCTCTTTTCAAGCTATCAAGTCTCTGGAAGATGACCCA	1992
Db	1922	AACAGAGAATCCACACTGTGTGCTCTTTTCAAGCTATCAAGTCTCTGGAAGATGACCCA	1981
QY	1993	ACTTGCAAAACAAAGGAATTTTGGGCCACCATTTTGGATTCCAGCAGGTCCTCTTCTGAC	2052
Db	1982	ACTTGCAAAACAAAGGAATTTTGGGCCACCATTTTGGATTCCAGCAGGTCCTCTTCTGAC	2041
QY	2053	TATACTCTCCAGAAATTTCTCTTTTGAAGTACTACTGGAATTTACATTCGTATGGGATG	2112
Db	2042	TATACTCTCCAGAAATTTCTCTTTTGAAGTACTACTGGAATTTACATTCGTATGGGATG	2101
QY	2113	CTCTACAGCCTCATGATCTACAGCCTGGAAGAAATATCTCTACTGTGCTGCTTCATATAT	2172
Db	2102	CTCTACAGCCTCATGATCTACAGCCTGGAAGAAATATCTCTACTGTGCTGCTTCATATAT	2161
QY	2173	GGTGGTCTCAGGTGCAAGTTGGTGAATTAATTCGGTTTAAAGGAGTCAAGTATTTCCGCTTG	2232
Db	2162	GGTGGTCTCAGGTGCAAGTTGGTGAATTAATTCGGTTTAAAGGAGTCAAGTATTTCCGCTTG	2221
QY	2233	AATACCTCTAGCTCTCTAGTTATGTGGTTGTAGTGATAGACAAACAGGGGATCCTGTAC	2292
Db	2222	AATACCTCTAGCTCTCTAGTTATGTGGTTGTAGTGATAGACAAACAGGGGATCCTGTAC	2281
QY	2293	CGAGGGCTTAAATTTGAAGCGGCTTTTAAATATAAATGGGTCAAATAGAAATTTGACGAT	2352
Db	2282	CGAGGGCTTAAATTTGAAGCGGCTTTTAAATATAAATGGGTCAAATAGAAATTTGACGAT	2341
QY	2353	CAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTTCATTTGACTTAGATCGTGTG	2412
Db	2342	CAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTTCATTTGACTTAGATCGTGTG	2401
QY	2413	GGCATCCACGGCTGGTCTTATGGAGGATACCTCTCCCTGATGGCATTAATCCAGAGTCA	2472
Db	2402	GGCATCCACGGCTGGTCTTATGGAGGATACCTCTCCCTGATGGCATTAATCCAGAGTCA	2461
QY	2473	GATATCTTTCAGGGTGTCTATTTGCTGGGGCCCGAGTCACCTCTGTGGATCTTCTATGATACA	2532
Db	2462	GATATCTTTCAGGGTGTCTATTTGCTGGGGCCCGAGTCACCTCTGTGGATCTTCTATGATACA	2521
QY	2533	GGATACACGGAACGTTTATATATGGGTCACTGACCAAGATGAACAGGGGCTATTTACTTAGGA	2592
Db	2522	GGATACACGGAACGTTTATATATGGGTCACTGACCAAGATGAACAGGGGCTATTTACTTAGGA	2581
QY	2593	TCTGTGGCCATGCAAGCAGAAAGTTTCCCTCTCTGAAACCAAAATCGTTTACTGCTTACAT	2652
Db	2582	TCTGTGGCCATGCAAGCAGAAAGTTTCCCTCTCTGAAACCAAAATCGTTTACTGCTTACAT	2641
QY	2653	GGTTTCTCGATGAGAAATGTCATTTTGGACATACAGTATATTTACTGAGTTTTTATGAGT	2712
Db	2642	GGTTTCTCGATGAGAAATGTCATTTTGGACATACAGTATATTTACTGAGTTTTTATGAGT	2701
QY	2713	AGGCTCGGAAGCCATATGATTTTACAGATCTATCTCTCAGGAGACACAGCATTAAGAGTT	2772
Db	2702	AGGCTCGGAAGCCATATGATTTTACAGATCTATCTCTCAGGAGACACAGCATTAAGAGTT	2761
QY	2773	CCTGAATCGGGAGAACATTTATGAACCTGCACTCTTTTGGCACTACCTTCAAGAAAACCTTGA	2832
Db	2762	CCTGAATCGGGAGAACATTTATGAACCTGCACTCTTTTGGCACTACCTTCAAGAAAACCTTGA	2821
QY	2833	TCAGTATTCCTGCTCTTAAAGTGATATAATTTTGCACCTGTGTAGAACTCTCTGGTATAC	2892
Db	2822	TCAGTATTCCTGCTCTTAAAGTGATATAATTTTGCACCTGTGTAGAACTCTCTGGTATAC	2881
QY	2893	ACTGGCTATTTAAACCAATGAGGAGGTTTAAATCAACAGAAAACACAGAAATTTGATCATCAC	2952
Db	2882	ACTGGCTATTTAAACCAATGAGGAGGTTTAAATCAACAGAAAACACAGAAATTTGATCATCAC	2941
QY	2953	ATTTTGATACCTGCCATGTAACATCTACTCTCCGAAATAAATTTGTGGTCCCATGAGGGGT	3012
Db	2942	ATTTTGATACCTGCCATGTAACATCTACTCTCCGAAATAAATTTGTGGTCCCATGAGGGGT	3001
QY	3013	CTACGGTTTCTGGTGTAGTAAATCTAAATACCTTAAACCCCATGCTCAAAATCAAATGATACA	3072
Db	3002	CTACGGTTTCTGGTGTAGTAAATCTAAATACCTTAAACCCCATGCTCAAAATCAAATGATACA	3061
QY	3073	TATTCCTGAGAGACCCAGCAATACCAATAAGAAATTTACTAAAAAAA 3117	
Db	3062	TATTCCTGAGAGACCCAGCAATACCAATAAGAAATTTACTAAAAAAA 3106	
RESULT 4			
LOCUS	AX354793	Sequence 1 from Patent WO0179473.	
DEFINITION	AX354793	3143 bp DNA linear PAT 06-FEB-2002	
ACCESSION	AX354793		
VERSION	AX354793.1	GI:18619526	
KEYWORDS			
SOURCE			
ORGANISM			
Homo sapiens (human)			
Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.			
1			
Meyers, R.A. and Williamson, M.			
21953, a human prolyl oligopeptidase family member and uses thereof			
Patent: WO 0179473-A 1 25-OCT-2001;			
Millennium Pharmaceuticals, Inc. (US)			
Location/Qualifiers			
1. 3143			
/organism="Homo sapiens"			
/mol_type="unassigned DNA"			

CDS	/db_xref="taxon:9606" 229..2877		/note="unnamed protein product"		/codon_start=1 /protein_id="CAD22990.1" /db_xref="GI:18619527"		/translation="MAAMETEQLGVIEFETADCEBIESODRPLEPFFVVERYSWSQ LKKLLADTRKHYGYMAKAPHDPMFVKRNDPDGPHSDRIYVYLAWSGENRENTLFSYSEI PKTINRAAVLMWSKPLDLFQATLDYGYMSREELLRRKRIGTVGLIASDYHQGG TFLFAGGSGIVHVQGGPQOPLRPNLIVETSCNRMDPKLCPADPDMWIAFHN DIWISNVTREERLLTVHNLMBEEDARSAGVATFVLQBEFDRISGYMCPKAEIT PSGGKILRLVEENDESEVEI IHVTSMPLETRADSFRYPKGTANPKVTPTEKMEIMI DAEGRIIDVIDKEILQPEILFEGVEYLARAGWTEGKYAWSIILLDRSOTRLOJLVLS PELFPVEDDWMERQRLIESVPSVPLLIYBETDWINHDIHFVPSHREEIEF IFASQKTFGRHLVKITSILKESYKRSGGLPAPSPDFKPIKEBIAITSGEWEVLGR HGSNTOQDEVRRLVYFEGTKDSPLLEHLYVYVSNPGEVTRITDRGYSHSCCISQHQD FFISKYSNOKPHCVSLYKLSPPEDPTCKTEFWATILDSAGPLPDYTPPEIEPSFS TTGFTLYGMLYKPHDLOPKKYPTVLFIYGGQVOLVNNRPFKGVKYLRLNTLASLGVV VVVDNRSGCHRGKLFKGAFFKMGQIEIDQVEGLQYLASRYDFIDLDVRVGHWSY GGYLSLMALMORSDFRVAJAGAPVTLWIFDYGTERTYRMYGHPDQNEGGYLYLGSVAMQ AEKFPSEPNRLLHGFLENVHFAHTSILLSPFVRAGKPYDLQIYPOERHSIRVPS GEHYELHLLHYLQENLGSRIAALKVI"		Query Match Best Local Similarity 94.8%; Score 2957; DB 6; Length 3143; Matches 3107; Conservative 0; Mismatches 3; Indels 0; Gaps 0;			
QY	11	GCCTCCGAGCCCAAGCCGCTGCTACTGCCGCCGCTCTTCTTTAGTGGCCGCTTCCGCCG 70										
Db	26	GCCTCCGAGCCCAAGCCGCTGCTACTGCCGCCGCTCTTCTTTAGTGGCCGCTTCCGCCG 85										
QY	71	CTGGGTGTGCACCGCGCGCGCGCGAGGAAGCCATGCAACAGGACCGGAGTGGAGGC 130										
Db	86	CTGGGTGTGCACCGCGCGCGCGCGAGAGGCCACTACAAACAGGACCGGAGTGGAGGC 145										
QY	131	GGCGCAGCATGAAGCGCGCGCGCGCTCATAGCGCAGCGTGGGAGCGGTCGGCGGG 190										
Db	146	GGCGCAGCATGAAGCGCGCGCGCGCTCCATAGCGCAGCGTGGGAGCGGTCGGCGGG 205										
QY	191	GCGGGGGGAAGAAATGCAACATGGCAGCAGCAATGGGAAACAGAACAGCTGGGTGTG 250										
Db	206	GCGGGGGGAAGAAATGCAACATGGCAGCAGCAATGGGAAACAGAACAGCTGGGTGTG 265										
QY	251	AGATATTTGAACTCGGAGCTGTGAGGAGATATTTGAATCACAGGATCGGCTAAATGG 310										
Db	266	AGATATTTGAACTCGGAGCTGTGAGGAGATATTTGAATCACAGGATCGGCTAAATGG 325										
QY	311	AdGCTTTTATGTTGAGCGGTATTCTCGAGTCAGCTTAAAGAGCTGCTTCCGATACCA 370										
Db	326	AdGCTTTTATGTTGAGCGGTATTCTCGAGTCAGCTTAAAGAGCTGCTTCCGATACCA 385										
QY	371	GAAATATCATGGCTACATGATGGCTAAGGCACCAATGATTTTCATGTTTGTGAAGAGA 430										
Db	386	GAAATATCATGGCTACATGATGGCTAAGGCACCAATGATTTTCATGTTTGTGAAGAGA 445										
QY	431	ATGATCCAGATGGAACCTCATTCAGACAGATCTATTAACCTTGGCCATGCTCTGGTAGACA 490										
Db	446	ATGATCCAGATGGAACCTCATTCAGACAGATCTATTAACCTTGGCCATGCTCTGGTAGACA 505										
QY	491	GAGAAATACATGTTTTATTTCTGAATTTCCAAAACCTATCAATAGCAGCAGCTCTTAA 550										
Db	506	GAGAAATACATGTTTTATTTCTGAATTTCCAAAACCTATCAATAGCAGCAGCTCTTAA 565										
QY	551	TGCTCTCTTGGAGGCTCTTTTGGATCTTTTTCAGGCAACACTGGACATGGAATGTAAT 610										
Db	566	TGCTCTCTTGGAGGCTCTTTTGGATCTTTTTCAGGCAACACTGGACATGGAATGTAAT 625										
QY	611	CTCGAGAGAGAACTATTTAAGAGAGAGAGAAACCACTTGGAAACAGTTCGGAATTTGCTTT 670										
Db	626	CTCGAGAGAGAGAACTATTTAAGAGAGAGAGAAACCACTTGGAAACAGTTCGGAATTTGCTTT 685										
QY	671	ACGATTTATCACCAAGAGAGTGGAAACATTTCTGTTTTCAGCGCGGTAGTGGAAATTTATCACG 730										

Db	1766	TCCAAGTTGATGAAGTCAGAAGCGCTGGTATATTTTGAAGGCACCAAGAAGCTCCCTCTTTAG	1822
Qy	1811	AGCATCACCTGTACGTAGTCAGTTTACGTAATCCTGGAGAGGTGACAAGGCTGACTGAC	1870
Db	1826	AGCATCACCTGTACGTAGTCAGTTTACGTAATCCTGGAGAGGTGACAAGGCTGACTGAC	1885
Qy	1871	GTGGTACTCACAATCTTGTCTGCATCAGTCAGCACGTGTGACTCTTTTATAAGTAAGTATA	1930
Db	1886	GTGGCTACTCACAATCTTGTCTGCATCAGTCAGCACGTGTGACTCTTTTATAAGTAAGTATA	1945
Qy	1931	GTAAACAGAAGAAATCCACACTGTGTGCCCTTTTACAAGCTATCAAAGTCCTCGAAGTAC	1990
Db	1946	GTAAACAGAAGAAATCCACACTGTGTGCCCTTTTACAAGCTATCAAAGTCCTCGAAGTAC	2005
Qy	1991	CAACTTGCAAAACAAAGGAAATTTTGGGCGCACCAATTTTGGATTACGACAGTCTCTCTCGT	2050
Db	2006	CAACTTGCAAAACAAAGGAAATTTTGGGCGCACCAATTTTGGATTACGACAGTCTCTCTCGT	2065
Qy	2051	ACTATATCTCTCCAGAAATTTCTCTTTTGARAGTACTACTGGATTACATTTGATATGGA	2111
Db	2066	ACTATATCTCTCCAGAAATTTCTCTTTTGARAGTACTACTGGATTACATTTGATATGGA	2125
Qy	2111	TGCTCTACAAGCCTCATGATCTACAGCCTGGAAGAAATATCCTACTGCTGCTTCATAT	2170
Db	2126	TGCTCTACAAGCCTCATGATCTACAGCCTGGAAGAAATATCCTACTGCTGCTTCATAT	2185
Qy	2171	ATGGTGGTCCTCAGTGCAGTGTGTGTAATTAATCGGTTTAAAGAGGTCAAGTATTTCCGCT	2230
Db	2186	ATGGTGGTCCTCAGTGCAGTGTGTGTAATTAATCGGTTTAAAGAGGTCAAGTATTTCCGCT	2245
Qy	2231	TGAATACCTTAGCCCTCTCTAGGTTATGTGTTGTAGTGATAGACAACAGGGGATCTGTCT	2290
Db	2246	TGAATACCTTAGCCCTCTCTAGGTTATGTGTTGTAGTGATAGACAACAGGGGATCTGTCT	2305
Qy	2291	ACCAGGGGCTTAAATTTGAAGCGCCCTTTAAATATATAAATGGGTCAATAGAAAATTGACG	2350
Db	2306	ACCAGGGGCTTAAATTTGAAGCGCCCTTTAAATATATAAATGGGTCAATAGAAAATTGACG	2365
Qy	2351	ATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTCAATGCACTTAGAATCGTG	2410
Db	2366	ATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTCAATGCACTTAGAATCGTG	2425
Qy	2411	TGGGATCCACAGGCTGGTCCCTATGGAGGATACCTCTCCCTGATGGCATTAAATGCAAGAGT	2470
Db	2426	TGGGATCCACAGGCTGGTCCCTATGGAGGATACCTCTCCCTGATGGCATTAAATGCAAGAGT	2485
Qy	2471	CAGATATCTTCAGGTTGCTATTCCTGGGCCCCCAGTCACTCTGTGGATCTTCTATGATA	2530
Db	2486	CAGATATCTTCAGGTTGCTATTCCTGGGCCCCCAGTCACTCTGTGGATCTTCTATGATA	2545
Qy	2531	CAGGATACACGNAAGTTATATGGGTCAACCTGACCAAGATGNAACAGGCTATTACTTAG	2590
Db	2546	CAGGATACACGNAAGTTATATGGGTCAACCTGACCAAGATGNAACAGGCTATTACTTAG	2605
Qy	2591	GATCTGTGGCCATGCAAGCAGAAAAGTTTCCCTCTGAAACCAATCGTTTACTGTCTCTAC	2650
Db	2606	GATCTGTGGCCATGCAAGCAGAAAAGTTTCCCTCTGAAACCAATCGTTTACTGTCTCTAC	2665
Qy	2651	ATGGTTTCTCTGGATGAGAAATGTCATTTTGCAATACCAAGTATATCTAGTGTTTTTTAG	2710
Db	2666	ATGGTTTCTCTGGATGAGAAATGTCATTTTGCAATACCAAGTATATCTAGTGTTTTTTAG	2725
Qy	2711	TGAGGGCTGGAAGCCATATGATTTTACAGATCTATCTCTCAGGAGAGACACAGCATAGAG	2770
Db	2726	TGAGGGCTGGAAGCCATATGATTTTACAGATCTATCTCTCAGGAGAGACACAGCATAGAG	2785
Qy	2771	TTCTCGAATCGGGAGAACATATGAATGCACTCTTTTGGCACTACCTTCAAGAAAACTTTG	2830
Db	2786	TTCTCGAATCGGGAGAACATATGAATGCACTCTTTTGGCACTACCTTCAAGAAAACTTTG	2845
Qy	2831	GATCACGATTTGCTGCTCTTAAAGTGATATAAATTTTGACCTGTGTAGAACTCTCTGGTAT	2890
Db	2846	GATCACGATTTGCTGCTCTTAAAGTGATATAAATTTTGACCTGTGTAGAACTCTCTGGTAT	2905

	Qy	2891	ACACTGGCTATTAAACCAATAGGAGGTTTAATCAACAGAAAAACAACAATTCATCATC	2995
	Dd	2906	ACACTGGCTATTAAACCAATAGGAGGTTTAATCAACAGAAAAACAACAATTCATCATC	2965
	Qy	2951	ACATTTTGATACCTCGGCATGTAAACATCTACTCTTGAATAAATGTGTGCCATGCAGGG	3010
	Dd	2966	ACATTTTGATACCTCGGCATGTAAACATCTACTCTTGAATAAATGTGTGCCATGCAGGG	3025
	Qy	3011	GCTACGGTTGTGGTAGTAATCTAATPACCTTAAACCCCAATGCTCAAATAAATGATA	3070
	Dd	3026	GCTACGGTTGTGGTAGTAATCTAATPACCTTAAACCCCAATGCTCAAATAAATGATA	3085
	Qy	3071	CATATTCCTGAGAGACCAGCAATACCATAGAATTTACTTAAAAAAA	3120
	Dd	3086	CATATTCCTGAGAGACCAGCAATACCATAGAATTTACTTAAAAAAA	3135
RESULT 5				
AR631283				
LOCUS				
DEFINITION				
ACCESSION				
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
FEATURES				
source				
ORIGIN				
Query Match				
Best Local Similarity				
Matches 319; Conservative				
Score 2889; DB 6; Length 4829;				
Pred. No. 0; Indels 2; Gaps 1;				
Qy	1	AAGTGTAAAGCCTCCGAGGCCAAGCGCTGTCTACTGCGCGCGTCTCTTAGTGGCG	60	
Dd	1	AAGTGTAAAGCCTCCGAGGCCAAGCGCTGTCTACTGCGCGCGTCTCTTAGTGGCG	60	
Qy	61	CGTTCGCGCCTCGGTTGTCAACGCGCGCGCGCGAGGAGCACTGCAACAGGACCG	120	
Dd	61	CGTTCGCGCCTCGGTTGTCAACGCGCGCGCGCGAGGAGCACTGCAACAGGACCG	120	
Qy	121	GAGTGGAGCGCGCAGCATGAAGCGCGGAGGCGCGCTCTCCTAGCGCACGTGGGACCG	180	
Dd	121	GAGTGGAGCGCGCAGCATGAAGCGCGGAGGCGCGCTCTCCTAGCGCACGTGGGACCG	180	
Qy	181	TCCGGCGCGGCGCGGGGGAGGAAAAATGCAACATGCGAGCAATGAAACAGAACAG	240	
Dd	181	TCCGGCGCGGCGCGGGGGAGGAAAAATGCAACATGCGAGCAATGAAACAGAACAG	240	
Qy	241	CTGGGTGTTGAGATATTTGAAACTCGGACGTGAGAGAGAAATTTGAATCAGAGATCG	300	
Dd	241	CTGGGTGTTGAGATATTTGAAACTCGGACGTGAGAGAGAAATTTGAATCAGAGATCG	300	
Qy	301	CCTAAATTCGAGCGCTTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAAAGCTGCTT	360	
Dd	301	CCTAAATTCGAGCGCTTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAAAGCTGCTT	360	
Qy	361	GCCGATACCAAGAAATATCATGGCTTACATGATGGCTAAGGCACACATGATTTCATGTTT	420	
Dd	361	GCCGATACCAAGAAATATCATGGCTTACATGATGGCTAAGGCACACATGATTTCATGTTT	420	
Qy	421	GTGAAGAGGAATGATCCAGATGGACCTCATTCAGACAGAAATCTATTACCTTGCCATGCT	480	

[illegible]

1501	TGGATAAATATCCATGCATCTTTTCATGTGTTTTTCCCAGAAAGTCACGAAGAGGAAATTGAG	1566
1561	TTTATTTTTTGGCTCTGAATGCAAAAAGAGTTTTCCGTCAATTTATACAAAAATTACATCTATT	1620
1561	TTTATTTTTTGGCTCTGAATGCAAAAAGAGTTTTCCGTCAATTTATACAAAAATTACATCTATT	1620
1621	TTAAAGGAAAGCAAAATATAACGATCCAGTGTGTGGCTGCCTGTCCAAGTGATTTCCAAG	1680
1621	TTAAAGGAAAGCAAAATATAACGATCCAGTGTGTGGCTGCCTGTCCAAGTGATTTCCAAG	1680
1681	TGTCCTTATCAAGAGGAGATAGCAATTAACGATGTGTGAATGGGAAGTTCTTGGCCGGCAT	1740
1681	TGTCCTTATCAAGAGGAGATAGCAATTAACGATGTGTGAATGGGAAGTTCTTGGCCGGCAT	1740
1741	GGATCTAATATCCCAAGTTGATGAAGTCAGAAAGCTGGTATATTTTTGAAGGCACCAAGAC	1800
1741	GGATCTAATATCCCAAGTTGATGAAGTCAGAAAGCTGGTATATTTTTGAAGGCACCAAGAC	1800
1801	TCCCTTTAGAGCATCACCTGTACGTAGTCAGTTACGTAAATCCTGGAGAGGTCACAAGG	1860
1801	TCCCTTTAGAGCATCACCTGTAGTAGTCAGTTACGTAAATCCTGGAGAGGTCACAAGG	1860
1861	CTGACTGACCGTGGCTACTCACATTTCTTGTGCATCAGTCAGCACTGTGACTCTTTTATA	1920
1861	CTGACTGACCGTGGCTACTCACATTTCTTGTGCATCAGTCAGCACTGTGACTCTTTTATA	1920
1921	AGTAAGTATAGTAAACAGAGAATCCACACTGTGTGTCCCTTTTACAAGCTATCAAGTCCT	1980
1921	AGTAAGTATAGTAAACAGAGAATCCACACTGTGTGTCCCTTTTACAAGCTATCAAGTCCT	1980
1981	GAAGATGACCCAACTTGCAGAAAACAAAGGAATTTTTGGGCCACCATTTTGGATTTACAGAGT	2040
1981	GAAGATGACCCAACTTGCAGAAAACAAAGGAATTTTTGGGCCACCATTTTGGATTTACAGAGT	2040
2041	CCTCTTCTGCTACTATCTCCTCAGAAAATTTTCTCTTTTGAAGTACTACTGGAATTTACA	2100
2041	CCTCTTCTGCTACTATCTCCTCAGAAAATTTTCTCTTTTGAAGTACTACTGGAATTTACA	2100
2101	TTGTATGGATGCTCTACAAGCCTCATGATCTACAGCCTGGAAGAAATATCCTACTGTG	2160
2101	TTGTATGGATGCTCTACAAGCCTCATGATCTACAGCCTGGAAGAAATATCCTACTGTG	2160
2161	CTGTTTCATATATGTGG--TCCTCAGGTACAGTTGGTGAATAATCGGTTTAAAGGAGTCA	2218
2161	CTGTTTCATATATGTGGTCTCCTCAGGTACAGTTGGTGAATAATCGGTTTAAAGGAGTCA	2220
2219	AGTATTTCCGCTTGAATACCCCTAGCCTCTCTAGTTTATGTGGTTGTAGTAGTAGACAACA	2278
2221	AGTATTTCCGCTTGAATACCCCTAGCCTCTCTAGTTTATGTGGTTGTAGTAGACAACA	2280
2279	GGGGATCCTGTGACCGAGGGCTTAAATTTGAAGGGCCCTTTAAATATATAAATGGTCAAA	2338
2281	GGGGATCCTGTGACCGAGGGCTTAAATTTGAAGGGCCCTTTAAATATATAAATGGGTCAAA	2340
2339	TAGAAATTTGACGATCAGGTGGAGGACTCAATATCTAGCTTCTCGATATGATTTCAATG	2398
2341	TAGAAATTTGACGATCAGGTGGAGGACTCCAATATCTAGCTTCTCGATATGATTTCAATG	2400
2399	ACTTAGATCTGTGGGATCCACGGCTGGTCTCTATGAGGAGTACTCTCTCCCTGATGGCAT	2458
2401	ACTTAGATCTGTGGGATCCACGGCTGGTCTCTATGAGGAGTACTCTCTCCCTGATGGCAT	2460
2459	TAATGCAGAGGTACAGATATCTTCAGGGTTGCTATTGCTGGGGCCCCAGTCACTCTGTGGA	2518
2461	TAATGCAGAGGTACAGATATCTTCAGGGTTGCTATTGCTGGGGCCCCAGTCACTCTGTGGA	2520
2519	TCTTCTTATGATACAGGATACACGGAAAGTTATATUGGGTCAACCTTGACAGATGAACAGG	2578
2521	TCTTCTTATGATACAGGATACACGGAAAGTTATATUGGGTCAACCTTGACAGATGAACAGG	2580
2579	GCTATTACTTAGGATCTGTGGCCATGCAAGGAGAAAGTTCCCTCTCTGAACCAATCGTT	2638
2581	GCTATTACTTAGGATCTGTGGCCATGCAAGGAGAAAGTTCCCTCTCTGAACCAATCGTT	2640

```
QY 2639 TACTGCTCTTACATGGTTTCTCGATGAGAAATGTCATTTGACATACACAGTATATTAC 2698
Db 2641 TACTGCTCTTACATGGTTTCTCGATGAGAAATGTCATTTGACATACACAGTATATTAC 2700
QY 2699 TGAGTCTTTTATGAGGGCTGGAAAGCCATATGATTTACAGATCTATCTCTCAGAGAGAC 2758
Db 2701 TGAGTCTTTTATGAGGGCTGGAAAGCCATATGATTTACAGATCTATCTCTCAGAGAGAC 2760
QY 2759 ACAGCTAAGAGTTCTGGAATCGGAGAAACATTTAGAACTGCACTTTTGGCACTTCTTC 2818
Db 2761 ACAGCTAAGAGTTCTGGAATCGGAGAAACATTTAGAACTGCACTTTTGGCACTTCTTC 2820
QY 2819 AAGAAAACCTTGGATCACGPTATTTGCTGCTCTAAAGTGTATAAATTTTGACCTGTGAGA 2878
Db 2821 AAGAAAACCTTGGATCACGPTATTTGCTGCTCTAAAGTGTATAAATTTTGACCTGTGAGA 2880
QY 2879 ACTCTCTGGTATACACTGGCTATTTAAACCAATGAGGAGGTTTAAATCAACAGAAAACACA 2938
Db 2881 ACTCTCTGGTATACACTGGCTATTTAAACCAATGAGGAGGTTTAAATCAACAGAAAACACA 2940
QY 2939 GAAATTGATCATCATTTTGTATACCTGCCATGTAACATCTACTCTGAAAATAAATGTGG 2998
Db 2941 GAAATTGATCATCATTTTGTATACCTGCCATGTAACATCTACTCTGAAAATAAATGTGG 3000
QY 2999 TGCCATCAGGGGTCTACGGTTTGTGTAGTAACTTAATACCTTAACCCACATGCTCAA 3058
Db 3001 TGCCATCAGGGGTCTACGGTTTGTGTAGTAACTTAATACCTTAACCCACATGCTCAA 3060
QY 3059 AATCAATGATACATATCTCTGAGAGCCAGCAATACCAATGAAATTAATCAAAAAAAA 3118
Db 3061 AATCAATGATACATATCTCTGAGAGCCAGCAATACCAATGAAATTAATCAAAAAAAA 3120
QY 3119 AA 3120
Db 3121 AA 3122

RESULT 6
AX608735 4829 bp DNA linear PAT 17-FEB-2003
LOCUS AX608735
DEFINITION Sequence 12 from Patent WO0231134.
ACCESSION AX608735
VERSION AX608735.1 GI:28404303
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
REFERENCE 1
AUTHORS Qi,S., Akineanya,K.O., Riviere,P.J. and Junien,J.L.
TITLE Novel serine protease genes related to dppiv
JOURNAL Patent: WO 0231134-A 12 18-APR-2002;
Ferring BV (NL)
FEATURES
    source
        1..4829
            /organism="Homo sapiens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9606"

ORIGIN
Query Match 92.6%; Score 2889; DB 6; Length 4829;
Best Local Similarity 99.9%; Pred No. 0;
Matches 3119; Conservative 0; Mismatches 1; Indels 2; Gaps 1;

QY 1 AAGTGTAAAGCTCCGAGGCCAAGCCGCTGCTACTGCGCGCGCTGCTCTTAGTGCCG 60
Db 1 AAGTGTAAAGCTCCGAGGCCAAGCCGCTGCTACTGCGCGCGCTGCTCTTAGTGCCG 60
QY 61 CGTTCGCGCGCTGGTGTGACGCGCGCGCGCGCGAGGAAGCCACTGCAACAGAGACG 120
Db 61 CGTTCGCGCGCTGGTGTGACGCGCGCGCGCGCGAGGAAGCCACTGCAACAGAGACG 120
```

```
QY 121 GAGTGGAGGGCGGCGGCGAGCATGAAGCCGCGCAGCGCGCTCCATAGCGCACGTCGCGGACGG 180
Db 121 GAGTGGAGGGCGGCGGCGAGCATGAAGCCGCGCAGCGCGCTCCATAGCGCACGTCGCGGACGG 180
QY 181 TCCGGGCGGGCGGCGGCGGAGAAATGCAACATGGCAGCAGCAATGGAACAGAAACAG 240
Db 181 TCCGGGCGGGCGGCGGCGGAGAAATGCAACATGGCAGCAGCAATGGAACAGAAACAG 240
QY 241 CTGGGTGTTGAGATATTTGAAACTGCGGACTGTGAGGAGAAATTTGAATCAAGGATCGG 300
Db 241 CTGGGTGTTGAGATATTTGAAACTGCGGACTGTGAGGAGAAATTTGAATCAAGGATCGG 300
QY 301 CTTAAATTTGGAGCGCTTTTATGTTGAGCGGTATTCCTCGAGTGCAGCTTAAAAAGCTGCTT 360
Db 301 CTTAAATTTGGAGCGCTTTTATGTTGAGCGGTATTCCTCGAGTGCAGCTTAAAAAGCTGCTT 360
QY 361 GCCGATACAGAAAAATATCATGGCTACATGATGGCTAAGGCCACCATGATTTTATGTTT 420
Db 361 GCCGATACAGAAAAATATCATGGCTACATGATGGCTAAGGCCACCATGATTTTATGTTT 420
QY 421 GTGAAGAGGAATGATCCAGATGGACCTCATTCAGACAGAACTCTATTACCTTGGCCATGCT 480
Db 421 GTGAAGAGGAATGATCCAGATGGACCTCATTCAGACAGAACTCTATTACCTTGGCCATGCT 480
QY 481 GGTGAGAACAGAGAAATACACTGTTTATCTGAAATTTCCCAAACTATCAATAGAGCA 540
Db 481 GGTGAGAACAGAGAAATACACTGTTTATCTGAAATTTCCCAAACTATCAATAGAGCA 540
QY 541 GCAGTCTTAATGCTCTCTTGGAGCGCTCTTTGGATCTTTTCAGGGAACACTGGACTAT 600
Db 541 GCAGTCTTAATGCTCTCTTGGAGCGCTCTTTGGATCTTTTCAGGGAACACTGGACTAT 600
QY 601 GGAATGATTTCTCGAGAGAGAAACATTTAAGAGAGAAACGCAATTTGGAACTGCGGA 660
Db 601 GGAATGATTTCTCGAGAGAGAAACATTTAAGAGAGAAACGCAATTTGGAACTGCGGA 660
QY 661 ATTGCTTTTACGATTTATCACCAAGAGTGAACATTTCTGTTTCAAGCGGTAGTGA 720
Db 661 ATTGCTTTTACGATTTATCACCAAGAGTGAACATTTCTGTTTCAAGCGGTAGTGA 720
QY 721 ATTATCACGTAAGAGATGAGGGCCACAGGATTTAGCAACAACTTTAAGGCCCAAT 780
Db 721 ATTATCACGTAAGAGATGAGGGCCACAGGATTTAGCAACAACTTTAAGGCCCAAT 780
QY 781 CTAGTGGAAACTAGTTGTCCCAACATACGATGGAATCCAAATTTATGCCGCTGATCCA 840
Db 781 CTAGTGGAAACTAGTTGTCCCAACATACGATGGAATCCAAATTTATGCCGCTGATCCA 840
QY 841 GACTGGATTGCTTTTATACATAGCAACGATATTTGGATATCTAACATCGTAACCCAGAGAA 900
Db 841 GACTGGATTGCTTTTATACATAGCAACGATATTTGGATATCTAACATCGTAACCCAGAGAA 900
QY 901 GAAAGGAGACTCACTTTATGTCACAATGAGCTAGCAACATGGAAGAGATGCCAGATCA 960
Db 901 GAAAGGAGACTCACTTTATGTCACAATGAGCTAGCAACATGGAAGAGATGCCAGATCA 960
QY 961 GCTGGAGTCCGTACCTTTCTTCCAAAGAAATTTGATAGATATTCCTGGCTATTGGTGG 1020
Db 961 GCTGGAGTCCGTACCTTTCTTCCAAAGAAATTTGATAGATATTCCTGGCTATTGGTGG 1020
QY 1021 TGTCGCAAAAGCTGAAACAACTCCAGTGGTGAATTTCTAGNAATCTATATGAAGAA 1080
Db 1021 TGTCGCAAAAGCTGAAACAACTCCAGTGGTGAATTTCTAGNAATCTATATGAAGAA 1080
QY 1081 AATGATGATCTGAGTGGAAATTTATCTGTATCATCCCTATGTTGGAACAGAGGG 1140
Db 1081 AATGATGATCTGAGTGGAAATTTATCTGTATCATCCCTATGTTGGAACAGAGGG 1140
QY 1141 GCAGATTCACTCGTTATCCTTAAACAGAGTACAGAAATCTTAAAGTCACTTTTAAAGATG 1200
Db 1141 GCAGATTCACTCGTTATCCTTAAACAGAGTACAGAAATCTTAAAGTCACTTTTAAAGATG 1200
```


Db 2276 TGGCAACGATATTTGGATATCTAACATCGTAAACGAGAGAAAGGAGACTCACTTATGT 2335
QY 921 GCACATCAGCTAGCCACATGGAGAGATGCCAGATCAGCTGGAGTCGCTACCTTTGT 980
Db 2336 GCACATCAGCTAGCCACATGGAGAGATGCCAGATCAGCTGGAGTCGCTACCTTTGT 2395
QY 981 TCTCCAAAGAAATTTGATAGATATCTGGCTATTGGTGGTGTCCAAAGAGCTGAAACAAC 1040
Db 2396 TCTCCAAAGAAATTTGATAGATATCTGGCTATTGGTGGTGTCCAAAGAGCTGAAACAAC 2455
QY 1041 TCCAGTGGTGGTAAATCTTAGAATTTATATAGAGAAATGATGAATCTGAGGTGGA 1100
Db 2456 TCCAGTGGTGGTAAATCTTAGAATTTATATAGAGAAATGATGAATCTGAGGTGGA 2515
QY 1101 AATTATTCATGTTACATCCCTATGTTGGAAACAAGAGGGCAGATTCATTTCCGTTATCC 1160
Db 2516 AATTATTCATGTTACATCCCTATGTTGGAAACAAGAGGGCAGATTCATTTCCGTTATCC 2575
QY 1161 TAAACAGGTACAGCAAAATCCTAAAGTCACTTTTAAGATGTCAGAAATAATGATTTGATGC 1220
Db 2576 TAAACAGGTACAGCAAAATCCTAAAGTCACTTTTAAGATGTCAGAAATAATGATTTGATGC 2635
QY 1221 TGAAGAGGATCATAGATGTCATAGATAGGAATTAATTCACCTTTTGATTTCTATT 1280
Db 2636 TGAAGAGGATCATAGATGTCATAGATAGGAATTAATTCACCTTTTGATTTCTATT 2695
QY 1281 TGAAGAGGATCAATATATGTCAGAGCTGGATGAGTCTCTGAGGAAATATGCTTGTC 1340
Db 2696 TGAAGAGGATCAATATATGTCAGAGCTGGATGAGTCTCTGAGGAAATATGCTTGTC 2755
QY 1341 CATCTACTAGATCGCTCCAGACTCGCCTACAGATAGTGTGTGATCTCACCTGAAATATT 1400
Db 2756 CATCTACTAGATCGCTCCAGACTCGCCTACAGATAGTGTGTGATCTCACCTGAAATATT 2815
QY 1401 TATCCAGTAGAAGATGATTTATGGAAGGCGAGAGACTCATTTGAGTCAGTGCCTGATTC 1460
Db 2816 TATCCAGTAGAAGATGATTTATGGAAGGCGAGAGACTCATTTGAGTCAGTGCCTGATTC 2875
QY 1461 TGTGACGCCATTAATTTATGAGAAACAACAGACATCTGGATAAATATCCATGACAT 1520
Db 2876 TGTGACGCCATTAATTTATGAGAAACAACAGACATCTGGATAAATATCCATGACAT 2935
QY 1521 CTTTCATGTTTTTCCCAAGTCACGAGAGGAAATTTGATTTATTTTGGCTCTGATG 1580
Db 2936 CTTTCATGTTTTTCCCAAGTCACGAGAGGAAATTTGATTTATTTTGGCTCTGATG 2995
QY 1581 CAAACACAGTTTTCCGTCATTTATACAAATTTACATCTATTTTAAAGGAAAGCAATATAA 1640
Db 2996 CAAACACAGTTTTCCGTCATTTATACAAATTTACATCTATTTTAAAGGAAAGCAATATAA 3055
QY 1641 AGATCCAGTGGTGGCTGCTGCTCCAGTGATTTTCAAGTGTCTTATCAAAGAGGAGAT 1700
Db 3056 AGATCCAGTGGTGGCTGCTGCTCCAGTGATTTTCAAGTGTCTTATCAAAGAGGAGAT 3115
QY 1701 AGCAATTAACAGTGGTGAATGGGAAGTTCTTGGCCGGCATGGATCTAATATCCAAAGTTGA 1760
Db 3116 AGCAATTAACAGTGGTGAATGGGAAGTTCTTGGCCGGCATGGATCTAATATCCAAAGTTGA 3175
QY 1761 TGAAGTCAGAAGCTGGTGTATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACCT 1820
Db 3176 TGAAGTCAGAAGCTGGTGTATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACCT 3235
QY 1821 GTACGTAGTCAAGTAAATCTCGAGAGGTGACAAGGCTGACTGGACCGTGGTACTC 1880
Db 3236 GTACGTAGTCAAGTAAATCTCGAGAGGTGACAAGGCTGACTGGACCGTGGTACTC 3295
QY 1881 ACATTCCTGCTGCATCAGTCAGCTGAGTCTTTTATAGTAAGTATAGTACCCAGAA 1940
Db 3296 ACATTCCTGCTGCATCAGTCAGCTGAGTCTTTTATAGTAAGTATAGTACCCAGAA 3355
QY 1941 GAATCCACACTGTGTGCTTTTACAGCTATCAAGTCTCTGAAGATGACCCAACTTGCAA 2000
Db 3356 GAATCCACACTGTGTGCTTTTACAGCTATCAAGTCTCTGAAGATGACCCAACTTGCAA 3415

QY 2001 AACAAAGGAAATTTTGGGCCACCAATTTTGAATTCAGCAGGTCTCTTCTTGACTATACTCC 2060
Db 3416 AACAAAGGAAATTTTGGGCCACCAATTTTGAATTCAGCAGGTCTCTTCTTGACTATACTCC 3475
QY 2061 TCCAGAAATTTTCTCTTTTGAAGTACTACTGGAATTTACATTTGTATGGAGTCTCTACAA 2120
Db 3476 TCCAGAAATTTTCTCTTTTGAAGTACTACTGGAATTTACATTTGTATGGAGTCTCTACAA 3535
QY 2121 GCCTCATGATCTACAGCTCGAAAGAAATATCTCTACTGCTGTTTCAATATATGGTGGTCC 2180
Db 3536 GCCTCATGATCTACAGCTCGAAAGAAATATCTCTACTGCTGTTTCAATATATGGTGGTCC 3595
QY 2181 TCAGGTGCAAGTGTGTAATTAATCGGTTTAAAGAGTCAAGTATTTCCCGTTGAATACCT 2240
Db 3596 TCAGGTGCAAGTGTGTAATTAATCGGTTTAAAGAGTCAAGTATTTCCCGTTGAATACCT 3655
QY 2241 AGCCTCTCTAGGTTATGTTGTTAGTATAGACACAGGGGATCTCTGTCACCGAGGCT 2300
Db 3656 AGCCTCTCTAGGTTATGTTGTTAGTATAGACACAGGGGATCTCTGTCACCGAGGCT 3715
QY 2301 TAAATTTGAAGCGCCTTTAAATATAAAATGGGTCAAAATAGAAATTTGACGATCAGGTGGA 2360
Db 3716 TAAATTTGAAGCGCCTTTAAATATAAAATGGGTCAAAATAGAAATTTGACGATCAGGTGGA 3775
QY 2361 AGGACTCCAATATCTAGCTTCTCGATATGATTTCAATTGATCTTAGATCTGTTGGGATCCA 2420
Db 3776 AGGACTCCAATATCTAGCTTCTCGATATGATTTCAATTGATCTTAGATCTGTTGGGATCCA 3835
QY 2421 CGGCTGGTCTTATGAGGATACCTCTCCCTGATGGCATTAATTAATGACAGGTCAAGATCTT 2480
Db 3836 CGGCTGGTCTTATGAGGATACCTCTCCCTGATGGCATTAATTAATGACAGGTCAAGATCTT 3895
QY 2481 CAGGGTGTCTATGCTGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACAC 2540
Db 3896 CAGGGTGTCTATGCTGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACAC 3955
QY 2541 GGAACGTTATATGGTCAACCTGACAGGATGACAGGCTTACTTAGGATCTGTGC 2600
Db 3956 GGAACGTTATATGGTCAACCTGACAGGATGACAGGCTTACTTAGGATCTGTGC 4015
QY 2601 CATGACAGCAAGAAAGTTCCCTCTGAACCAATCGTTTACTGCTCTTACATGGTTCTCT 2660
Db 4016 CATGACAGCAAGAAAGTTCCCTCTGAACCAATCGTTTACTGCTCTTACATGGTTCTCT 4075
QY 2661 GGATGAGATGCTCAATTTTGCATACCATAGTATATTTAGTGTGAGGGCTGG 4135
Db 4076 GGATGAGATGCTCAATTTTGCATACCATAGTATATTTAGTGTGAGGGCTGG 4195
QY 2721 AAAGCCATATGATTTACAGATCTATCTCAGGAGAGACACAGCATTAAGAGTTCTGAATC 2780
Db 4136 AAAGCCATATGATTTACAGATCTATCTCAGGAGAGACACAGCATTAAGAGTTCTGAATC 4195
QY 2781 GGGAGACATTAATGAACTGCACTTTTGGACTACTTCAAGAAACCTTGGATCAGGTAT 2840
Db 4196 GGGAGACATTAATGAACTGCACTTTTGGACTACTTCAAGAAACCTTGGATCAGGTAT 4255
QY 2841 TGCTGCTCTAAAGTGTATATAATTTTGAACCTGTGTAGAACTCTCTGGTATACACTGGCTA 2900
Db 4256 TGCTGCTCTAAAGTGTATATAATTTTGAACCTGTGTAGAACTCTCTGGTATACACTGGCTA 4315
QY 2901 TTTTAACCAATAGGAGGTTTAAATCAACAGAAACACAGAAATTTGATTCATCATTTTGAAT 2960
Db 4316 TTTTAACCAATAGGAGGTTTAAATCAACAGAAACACAGAAATTTGATTCATCATTTTGAAT 4375
QY 2961 ACCTGCCATGTAACATCTACTCTCTGAAAATAAATTTGGTGCCTATGCAGGGGTCTACGGTT 3020
Db 4376 ACCTGCCATGTAACATCTACTCTCTGAAAATAAATTTGGTGCCTATGCAGGGGTCTACGGTT 4435
QY 3021 TGTGGTAGTAATCTTAATACCTTAAACCCCATGCTCAAAATCAAAATGATATATTTCTG 3080
Db 4436 TGTGGTAGTAATCTTAATACCTTAAACCCCATGCTCAAAATCAAAATGATATATTTCTG 4495

QY	3081	AGAGACCCAGCATATACCATAGAAATTACTATAAAAAAAAAAAAAA	3120
Db	4496	AGAGACCCAGCATATACCATAGAAATTACTATAAAAAAAAAAAAAA	4535
RESULT 8			
LOCUS	AY354202	3030 bp	mRNA linear PRI 02-SEP-2003
DEFINITION	Homo sapiens dipeptidyl peptidase 8 isoform 3 (DPP8) mRNA, complete cds, alternatively spliced.		
ACCESSION	AY354202		
VERSION	AY354202.1	GI:34329351	
KEYWORDS	.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE	1 (bases 1 to 3030)		
AUTHORS	Sha, J. H., Zhou, Z. M. and Li, J. M.		
TITLE	Direct Submission		
JOURNAL	Submitted (29-JUL-2003) Key Lab of Reproductive Medicine, Nanjing Medical University, 140 Han Zhong Road, Nanjing, Jiangsu 210029, China		
FEATURES	Location/Qualifiers		
source	1..3030		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/chromosome="15"		
	/tissue_type="testis"		
gene	1..3030		
	/gene="DPP8"		
CDS	161..2857		
	/gene="DPP8"		
	/note="alternatively spliced"		
	/codon_start=1		
	/product="dipeptidyl peptidase 8 isoform 3"		
	/protein_id="AAQ63887.1"		
	/db_xref="GI:34329352"		
	/translation="MKRSEQMKIKSGKNMAAMETQLGVEIPETADCBENIESQD RPKEFPFVERYSQKLLADTRYKHGYMMAKAPHDFMFVKRNDPDGPHSDRIYYL AMSGENRNTLFYSEIPKTIINRAAVLMSWPLDLFOALDLYDMYREBELLRERKR IGTVGIASVDYHGGSGTFLQAGSGIYHVKGQPGQFTQQPLRNLVETSCPNRMDDP KLCPADPDWIAFTHSNDIWSINVTREERLLTYVHNELANNEEDARAGVATFLVQBE FDYSGWPCPKAETTPSGKILRLIYEENDESEVEIITHVSPMLETRRADSPRYKTE GTANPKVTFKMSIMIDAEGRIDVIDKELIOPPEILFEGVEIYARAGMTPEGKYAWS ILLDRSOTRLOIIVLISPELFIPEVDVMERORLIESVPDSVTPLIIVEETTDIWINIH DIHFVPOQHEEBIEFTFASECKTGRFRLYKITSILKESKYRSGSLPAPDFKCPFI KEEIAITSGMEVLRHGSNIQVDSVRRLVFEGTKDSPLEHLYVVSYPNPEVTRL TRDYSKSCISIOHCDFFISKYSNKNPHCVSLYKLSPEDDPTCKTKEFWATILDSA GPLPDVTPPEIFSEFTGTGLYMLYKPHDLQPKGYPTVLTLYGGPQVQLVNNRFK GVKYFRLNTLASLYVYVVDNRGSRCHRLKFEAGFKYKMSQIISIDDPQVGLQYLAR YDFIDRLVGIHGWSYGGVLSLMALNQRSDIFRVALGAPVTLWIFDYDTGYTRYNGH PDONEQGYLGVSAMQAEKFPSEPNRLILLHGLFENVHFANTSILLSFLVRAGKPYD LQIYPQERHSIRVPESGEHYELHLHYLOENLGSRIAALKVI"		
ORIGIN			
Query Match 88.3%; Score 2756; DB 8; Length 3030;			
Best Local Similarity 99.9%; Pred. No. 0;			
Matches 2806; Conservative 0; Mismatches 1; Indels 0; Gaps 0;			
QY	201	AGGAAATGCAACATGCGACGACCAATGGAAACAGAACAGCTGGGTGTGAGATATTGA	260
Db	196	AGGAAATGCAACATGCGACGACCAATGGAAACAGAACAGCTGGGTGTGAGATATTGA	255
QY	261	AACTGCGGACTGTGAGGAGAAATATTGAATCACAGGATCGGCTAAATTTGAGCCTTTTA	320
Db	256	AACTGCGGACTGTGAGGAGAAATATTGAATCACAGGATCGGCTAAATTTGAGCCTTTTA	315
QY	321	TGTTGAGCGGTATTCTCTGGAGTCAGCTTAAAAAGCTGCTTGGCGATACCAAGAAATATCA	380
Db	316	TGTTGAGCGGTATTCTCTGGAGTCAGCTTAAAAAGCTGCTTGGCGATACCAAGAAATATCA	375

QY 1461 TGTGACGCCACTAATTTATCTATGAAGAAAACAACAGACATCTGGATAAATATCCATGACAT 1520
Db 1456 TGTGACGCCACTAATTTATCTATGAAGAAAACAACAGACATCTGGATAAATATCCATGACAT 1515
QY 1521 CTTTCATGTTTTTCCCAAGTCACGAAGAGGAATTCAGTTTATTTTTTGGCTCTGAAATG 1580
Db 1516 CTTTCATGTTTTTCCCAAGTCACGAAGAGGAATTCAGTTTATTTTTTGGCTCTGAAATG 1575
QY 1581 CAATAACAGGTTTCGGTCATTTATACAAAATTTACATCTATTTTAAAGGAAACCAATATAA 1640
Db 1576 CAATAACAGGTTTCGGTCATTTATACAAAATTTACATCTATTTTAAAGGAAACCAATATAA 1635
QY 1641 ACGATCCAGTGGTGGCTGCTCTCCAAAGTGATTTCAAGTGTCTCTATCAAAAGAGGAGAT 1700
Db 1636 ACGATCCAGTGGTGGCTGCTCTCCAAAGTGATTTCAAGTGTCTCTATCAAAAGAGGAGAT 1695
QY 1701 AGCAATTTACCAGTGGTGAATGGGAAGTTCTTGGCCGGCATGGATCTAATATCCAAAGTTGA 1760
Db 1696 AGCAATTTACCAGTGGTGAATGGGAAGTTCTTGGCCGGCATGGATCTAATATCCAAAGTTGA 1755
QY 1761 TGAAGTCAGAGGCTGGTATATTTTGAAGGCAACCAAGACTCCCTTTAGAGCATCACCT 1820
Db 1756 TGAAGTCAGAGGCTGGTATATTTTGAAGGCAACCAAGACTCCCTTTAGAGCATCACCT 1815
QY 1821 GTACGTAGTCAGTTACGTAAATCCTCGAGAGGTGCAAGGCTGACTGACCGTGGCTACTC 1880
Db 1816 GTACGTAGTCAGTTACGTAAATCCTCGAGAGGTGCAAGGCTGACTGACCGTGGCTACTC 1875
QY 1881 ACATTTCTGTGCATCAGTCAGCACGTGCATCTTTTATTAAGTAAGTATAGTACCAAGAA 1940
Db 1876 ACATTTCTGTGCATCAGTCAGCACGTGCATCTTTTATTAAGTAAGTATAGTACCAAGAA 1935
QY 1941 GAATCCACACTGTGTGCCCTTTTACAAAGCTATCAAGTCTCAAGATGACCCAACTTGCAA 2000
Db 1936 GAATCCACACTGTGTGCCCTTTTACAAAGCTATCAAGTCTCAAGATGACCCAACTTGCAA 1995
QY 2001 AACAAAGGAATTTTGGGCCACCAATTTTGAATTCAGCAGGTCCTCTTCCTGACTATATCC 2060
Db 1996 AACAAAGGAATTTTGGGCCACCAATTTTGAATTCAGCAGGTCCTCTTCCTGACTATATCC 2055
QY 2061 TCAGAAATTTTCTCTTTTGAAGTACTACTGATTTACATTTGATGGATGCTCTACAA 2120
Db 2056 TCAGAAATTTTCTCTTTTGAAGTACTACTGATTTACATTTGATGGATGCTCTACAA 2115
QY 2121 GCCTCATGATCTACAGGCTGGAAAGAAATATCCTACTGTCTGTTCATATATGCTGTCC 2180
Db 2116 GCCTCATGATCTACAGGCTGGAAAGAAATATCCTACTGTCTGTTCATATATGCTGTCC 2175
QY 2181 TCAGGTCAGTTGGTGAATAATCGGTTTAAAGGAGTCAAGTATTTTCGCTTGAATACCCCT 2240
Db 2176 TCAGGTCAGTTGGTGAATAATCGGTTTAAAGGAGTCAAGTATTTTCGCTTGAATACCCCT 2235
QY 2241 AGCCTCTCAGGTTATGCTGTGTAGTATAGACACAGGGGATCCTGTCAACGAGGCT 2300
Db 2236 AGCCTCTCAGGTTATGCTGTGTAGTATAGACACAGGGGATCCTGTCAACGAGGCT 2295
QY 2301 TAAATTTGAAGGCGCCTTTAAATATAAAATAGGGTCAAAATAGCAATTCAGGTTGA 2360
Db 2296 TAAATTTGAAGGCGCCTTTAAATATAAAATAGGGTCAAAATAGCAATTCAGGTTGA 2355
QY 2361 AGGACTCCAATATCTAGCTTCTCGATATGATTTCAATTTGATCTGTGGGATCCCA 2420
Db 2356 AGGACTCCAATATCTAGCTTCTCGATATGATTTCAATTTGATCTGTGGGATCCCA 2415
QY 2421 CAGCTGTCTATGAGGATACCTCTCCCTGATGGCATTTAATGACAGGTCAGATATCTT 2480
Db 2416 CAGCTGTCTATGAGGATACCTCTCCCTGATGGCATTTAATGACAGGTCAGATATCTT 2475
QY 2481 CAGGGTTGCTATTTGCTGGGCCCCAGTCACTCTGTGGATCTCTATGATACAGGATACAC 2540
Db 2476 CAGGGTTGCTATTTGCTGGGCCCCAGTCACTCTGTGGATCTCTATGATACAGGATACAC 2535
QY 2541 GGAACGTTATATGGGTCACCTCGACCAAGTGAACAGGGCTATTACTTTAGGATCTGTGGC 2600

Db 2536 GGAACGTTATATGGGTCACCTGACAGAAATGAACAGGGCTATTACTTTAGGATCTGTGGC 2595
QY 2601 CATCAACAGCAGAAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGTTTCTCT 2660
Db 2596 CATCAACAGCAGAAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGTTTCTCT 2655
QY 2661 GGATGAGAATGTCCATTTTGCACATACCAGTATATTACTGAGTATTTTGTAGTGGGCTGG 2720
Db 2656 GGATGAGAATGTCCATTTTGCACATACCAGTATATTACTGAGTATTTTGTAGTGGGCTGG 2715
QY 2721 AAAGCCATATGATTTACAGATCTATCTCAGGAGAGACACAGCATAAGAGTTCTGTAATC 2780
Db 2716 AAAGCCATATGATTTTACAGATCTATCTCAGGAGAGACACAGCATAAGAGTTCTGTAATC 2775
QY 2781 GGGAGAACATATGAACCTGATCTTTTGCACCTACCTTCAAGAAAACCTTGGATCACGAT 2840
Db 2776 GGGAGAACATATGAACCTGATCTTTTGCACCTACCTTCAAGAAAACCTTGGATCACGAT 2835
QY 2841 TGCTGCTCTAAAAGTGATATAATTTTGAACCTGTGTAGAACTCTCTGGTATACACTGGCTA 2900
Db 2836 TGCTGCTCTAAAAGTGATATAATTTTGAACCTGTGTAGAACTCTCTGGTATACACTGGCTA 2895
QY 2901 TTTTAAACCAATAGGAGGTTTAAATCAACAGAAAACACAGAAATGATCATCATTTTGCAT 2960
Db 2896 TTTTAAACCAATAGGAGGTTTAAATCAACAGAAAACACAGAAATGATCATCATTTTGCAT 2955
QY 2961 ACCTGCCATGTAAACATCTACTCTCTGAAAATAAATGTGTGCCATGCA 3007
Db 2956 ACCTGCCATGTAAACATCTACTCTCTGAAAATAAATGTGTGCCATGCA 3002

RESULT 9
AX354795 2649 bp DNA linear PAT 06-FEB-2002
LOCUS Sequence 3 from Patent WO0179473.
DEFINITION AX354795
ACCESSION AX354795
VERSION AX354795.1 GI:18619528
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE
1
Meyers, R.A. and Williamson, M.
21953, a human prolyl oligopeptidase family member and uses thereof
Patent: WO 0179473-A 3 25-OCT-2001;
Millennium Pharmaceuticals, Inc. (US)

FEATURES
source
1. .2649
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 81.6%; Score 2547; DB 6; Length 2649;
Best Local Similarity 99.9%; Pred. No; 0;
Matches 2647; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 214 ATGGCAGCAGCAATGGAACACAGAACAGCTGGGTGTGAGATATTTGAACCTGCGGACTGT 273
Db 1 ATGGCAGCAGCAATGGAACACAGAACAGCTGGGTGTGAGATATTTGAACCTGCGGACTGT 60
QY 274 GAGGAGATATTGGAATCAGAGGATCGGCTTAAATTTGAGGCTTTTATGTTGAGCGGTAT 333
Db 61 GAGGAGATATTGGAATCAGAGGATCGGCTTAAATTTGAGGCTTTTATGTTGAGCGGTAT 120
QY 334 TCCTGGAGTCAGCTTTAAAGAGCTGTGCGATACCAAGAAAATATCATGGCTTACATGATG 393
Db 121 TCCTGGAGTCAGCTTTAAAGAGCTGTGCGATACCAAGAAAATATCATGGCTTACATGATG 180
QY 394 GCTAAGGCCACCATGATTTTTCATGTTGTGAAGAGGAATGATCCAGATGGACCTCATTTCA 453

181 GCTAAGGACCAACATGATTTCAATGTTTGTGAAGGAATGATCCAGATGGACCTCATTTCA 240
454 GACAGAACTATTACCTTGCATGCTGTGTGAGAACAGAGAAATACACTGTTTATTTCT 513
241 GACAGAACTATTACCTTGCATGCTGTGTGAGAACAGAGAAATACACTGTTTATTTCT 300
514 GAAATTTCCAAATATCAATAGAGCAGCAGCTTTAATGCTCTCTTGGAAGCCCTTTTG 573
301 GAAATTTCCAAATATCAATAGAGCAGCAGCTTTAATGCTCTCTTGGAAGCCCTTTTG 360
574 GATCTTTTTCAGGCAACACTGGACTATGGAATGTATTTCTCGAGAGAAAGAACTATTAA 633
361 GATCTTTTTCAGGCAACACTGGACTATGGAATGTATTTCTCGAGAGAAAGAACTATTAA 420
634 GAAAGAAAACGCAATTTGGAACAGTCGGAATTTGCTTTACGATTTATCACCAAGGAAGTGA 693
421 GAAAGAAAACGCAATTTGGAACAGTCGGAATTTGCTTTACGATTTATCACCAAGGAAGTGA 480
694 ACATTTCTGTTTCAAGCCGGTAGTGGAAATTTATACGTTAAAGATGGAGGGCCACAAGGA 753
481 ACATTTCTGTTTCAAGCCGGTAGTGGAAATTTATACGTTAAAGATGGAGGGCCACAAGGA 540
754 TTTACGCAACAACTTTAAGGCCCAATCTAGTGGAAACTAGTTGTCCTCAACATACGGATG 813
541 TTTACGCAACAACTTTAAGGCCCAATCTAGTGGAAACTAGTTGTCCTCAACATACGGATG 600
814 GATCCAAAATTTATGCCCTGCTGATCCAGACTGGATTTGCTTTTATACATAGCAACGATTT 873
601 GATCCAAAATTTATGCCCTGCTGATCCAGACTGGATTTGCTTTTATACATAGCAACGATTT 660
874 TGSATATCTAACATCTGTACCAAGAGAGAAAGAGACTCACTTATATGTCACATAGACTA 933
661 TGSATATCTAACATCTGTACCAAGAGAGAAAGAGACTCACTTATATGTCACATAGACTA 720
934 GCCAACTATGGAAGAGATGCGAGATCAGCTGAGTGCCTACCTTTGTTTCTCCAAAGAA 993
721 GCCAACTATGGAAGAGATGCGAGATCAGCTGAGTGCCTACCTTTGTTTCTCCAAAGAA 780
994 TTTGATAGATATTCTGGCTATTGTTGGTGTCCAAAGCTGAAACAACTCCACAGTGGTGT 1053
781 TTTGATAGATATTCTGGCTATTGTTGGTGTCCAAAGCTGAAACAACTCCACAGTGGTGT 840
1054 AAAATCTTTAGAAATTTCTATATGAAGAAAATGATGAATCTGAGGTGGAAATTTATTCATGTT 1113
841 AAAATCTTTAGAAATTTCTATATGAAGAAAATGATGAATCTGAGGTGGAAATTTATTCATGTT 900
1114 ACATCCCTATGTTGGAACCAAGGGGAGATTCATTTCCGTTATCTTAAACAGGTACA 1173
901 ACATCCCTATGTTGGAACCAAGGGGAGATTCATTTCCGTTATCTTAAACAGGTACA 960
1174 GCAATCCTAAAGTCACTTTTAAGATGTCAGAAAATAATGATTTGATGCTGAAGGAGATC 1233
961 GCAATCCTAAAGTCACTTTTAAGATGTCAGAAAATAATGATTTGATGCTGAAGGAGATC 1020
1234 ATAGATGCTATAGATAAGNACTAATTCACACCTTTTGAGATTTCTATTGAGGAGTTGAA 1293
1021 ATAGATGCTATAGATAAGNACTAATTCACACCTTTTGAGATTTCTATTGAGGAGTTGAA 1080
1294 TATATTGCCAGAGCTGGATGGACTCTGAGGGAATAATGCTTGGTCCATCTCTACTAGAT 1353
1081 TATATTGCCAGAGCTGGATGGACTCTGAGGGAATAATGCTTGGTCCATCTCTACTAGAT 1140
1354 CGCTCCAGACTCGCTACAGATAGTTGATCTCACCTGAAATTTATTTATCCAGTAGAA 1413
1141 CGCTCCAGACTCGCTCGAGATAGTTGATCTCACCTGAAATTTATTTATCCAGTAGAA 1200
1414 GATGATGTTATGAAAGGAGAGACTCATTTGAGTCAGTGCCTGATTTCTGTAGCCCACTA 1473
1201 GATGATGTTATGAAAGGAGAGACTCATTTGAGTCAGTGCCTGATTTCTGTAGCCCACTA 1260
1474 ATTATCTATGAAGAAAACACAGACATCTGGATTAATATCCATGATCTTTTCATGTTTTT 1533

1261 ATTATCTATGAAGAAAACACAGACATCTCGATATAAATATCCATGACATCTTTCTATGTTTTT 1320
1534 CCCAAAGTCAACAGAGGAAATTTGAGTTTATTTTTGCTCTGAAATGCAAAACAGGTTTTT 1593
1321 CCCAAAGTCAACAGAGGAAATTTGAGTTTATTTTTGCTCTGAAATGCAAAACAGGTTTTT 1380
1594 CGTCATTTATACAAAATTTACATCTATTTTAAAGGAAACAAATATAAAACGATCCAGTGGT 1653
1381 CGTCATTTATACAAAATTTACATCTATTTTAAAGGAAACAAATATAAAACGATCCAGTGGT 1440
1654 GGGCTGCTCTCCAAAGTGTCTCAATATCTCAATATCTCAATATCTCAATATCTCAATATCTCA 1713
1441 GGGCTGCTCTCCAAAGTGTCTCAATATCTCAATATCTCAATATCTCAATATCTCAATATCTCA 1500
1714 GGTGAATGGGAAGTTCTTTGGCCGGCATGGATCTAATATCTCAAGTTGATGAAGTCAGAAAG 1773
1501 GGTGAATGGGAAGTTCTTTGGCCGGCATGGATCTAATATCTCAAGTTGATGAAGTCAGAAAG 1560
1774 CTGGTATATTTTGAAGGCCACCAAGACTCCCTCTTTAGAGCATCACCTCTGATAGTCAAGT 1833
1561 CTGGTATATTTTGAAGGCCACCAAGACTCCCTCTTTAGAGCATCACCTCTGATAGTCAAGT 1620
1834 TACGTAATCTCGAGAGGTGACAGGCTGACCTGACCGTGGCTACTACATTTCTTGCTGC 1893
1621 TACGTAATCTCGAGAGGTGACAGGCTGACCTGACCGTGGCTACTACATTTCTTGCTGC 1680
1894 ATCAGTCAGCACTGCTGACCTCTTTATAGTAAGTATAGTAAAGTATAGTAAAGTATAGTAA 1953
1681 ATCAGTCAGCACTGCTGACCTCTTTATAGTAAGTATAGTAAAGTATAGTAAAGTATAGTAA 1740
1954 GTGTCCCTTTACAAAGCTATCAAGTCTCTGAAAGTATGACCACTTGCGAAACAAAGGAATTT 2013
1741 GTGTCCCTTTACAAAGCTATCAAGTCTCTGAAAGTATGACCACTTGCGAAACAAAGGAATTT 1800
2014 TGGGCCACCAATTTTGAATTCAGCAGGTCTCTCTCTGACTATATCTCTCTCTCTCTCTCTCT 2073
1801 TGGGCCACCAATTTTGAATTCAGCAGGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1860
2074 TCTTTTGAAGTACTACTCGATTTTACATTTGATTTGGATGCTCTCAAGGCTCATGATCTA 2133
1861 TCTTTTGAAGTACTACTCGATTTTACATTTGATTTGGATGCTCTCAAGGCTCATGATCTA 1920
2134 CAGCTTGGAAAGAAATATCT 2193
1921 CAGCTTGGAAAGAAATATCT 1980
2194 GTGAATAATCCGTTTAAAGGAGTCAAGTATTTCCGCTTTGAATACCTTAGCCTCTCTAGGT 2253
1981 GTGAATAATCCGTTTAAAGGAGTCAAGTATTTCCGCTTTGAATACCTTAGCCTCTCTAGGT 2040
2254 TATGTGTTTCTAGTATAGACAAACAGGGGATCTCTGTCACCGAGGCTTTAAATTTGAAGGC 2313
2041 TATGTGTTTCTAGTATAGACAAACAGGGGATCTCTGTCACCGAGGCTTTAAATTTGAAGGC 2100
2314 GCCTTTAAATATATAAATGGTCAAAATAGAAAATGACGATCAGGTGGAGGACTCCAATAT 2373
2101 GCCTTTAAATATATAAATGGTCAAAATAGAAAATGACGATCAGGTGGAGGACTCCAATAT 2160
2374 CTAGCTTCTCGATATGATTTCACTGACTAGATCTGTTGGGATCTCCAGGCTGGCTCTAT 2433
2161 CTAGCTTCTCGATATGATTTCACTGACTAGATCTGTTGGGATCTCCAGGCTGGCTCTAT 2220
2434 GAGGATACCTCTCCCTGATGGCATTAATCAGAGGTGAGATATCTTCAGGGTTCGTATT 2493
2221 GAGGATACCTCTCCCTGATGGCATTAATCAGAGGTGAGATATCTTCAGGGTTCGTATT 2280
2494 GCTGGGGCCCCAGTCACTCTGTGGATCTCTATGATACAGGATACAGGAACTGTTATATG 2553
2281 GCTGGGGCCCCAGTCACTCTGTGGATCTCTATGATACAGGATACAGGAACTGTTATATG 2340
2554 GGTCACTCTGACAGAAATGAACAGGGCTATTACTTAGGATCTGTGGCATCTGCAAGCAGAA 2613
2341 GGTCACTCTGACAGAAATGAACAGGGCTATTACTTAGGATCTGTGGCATCTGCAAGCAGAA 2400

Qy	2614	AGTTTCCCTCTGAACCAAAATCGTTTACTGCTCTTACATGGTTTCTCTGGATGAGAATGTC	2673
Db	2401	AAAGTTCCCTCTGAACCAAAATCGTTTACTGCTCTTACATGGTTTCTCTGGATGAGAATGTC	2460
Qy	2674	CATTTTGCACATACCCAGTATATTACTGAGTTTCTTTTGTAGTCGGGCTGGAAAGCCATATGAT	2733
Db	2461	CATTTTGCACATACCCAGTATATTACTGAGTTTCTTTTGTAGTCGGGCTGGAAAGCCATATGAT	2520
Qy	2734	TTACAGATCTATCTCTCAGGAGAGACACAGCATAAAGATGTTCTCTGAATCCGGAGAACATTTAT	2793
Db	2521	TTACAGATCTATCTCTCAGGAGAGACACAGCATAAAGATGTTCTCTGAATCCGGAGAACATTTAT	2580
Qy	2794	GAATGCGATCTTTTGGCACTACCTTCAAGAAAAACCTTGGATCACGTATTGCTGCTCTAAAA	2853
Db	2581	GAATGCGATCTTTTGGCACTACCTTCAAGAAAAACCTTGGATCACGTATTGCTGCTCTAAAA	2640
Qy	2854	GTGATATAA	2862
Db	2641	GTGATATAA	2649
RESULT 10			
LOCUS	AY172659	2656 bp	mRNA linear PRI 08-JAN-2003
DEFINITION	Homo sapiens dipeptidyl peptidase IV-related protein-1 (DPRP1)		
ACCESSION	AY172659		
VERSION	AY172659.1	GI:27549549	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominoidea; Homo.		
REFERENCE	1 (bases 1 to 2656)		
AUTHORS	Qi, J., Akinsanya, K., Riviere, P. and Junien, J.-L.		
TITLE	Novel Serine Protease Genes Related To DPPIV		
JOURNAL	Patent: US (WO 0231134)-A 18-APR-2002;		
REFERENCE	2 (bases 1 to 2656)		
AUTHORS	Qi, J., Akinsanya, K., Riviere, P. and Junien, J.-L.		
TITLE	Direct Submission		
JOURNAL	Submitted (04-NOV-2002) Ferring Research Institute, 3550 General Atomics Ct., San Diego, CA 92121, USA		
FEATURES	Location/Qualifiers		
source	1..2656		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/chromosome="15"		
	/map="15q22.1-q22.2"		
gene	1..2656		
	/gene="DPRP1"		
CDS	8..2656		
	/gene="DPRP1"		
	/codon_start=1		
	/product="dipeptidyl peptidase IV-related protein-1"		
	/protein_id="AA017261.1"		
	/db_xref="GI:27549550"		
	/translation="MAAMETEQLGVIEIETADCENIETSDRPLEKLPFFVVERYSNSQ LKKLADTRKTHGYMAKAPDFMVRNDPDGSHDRIIYILAMSGNRENTILFYSEI PTLNRAALMLSKPLDLDFQATLDLYMYSEELRERKRIGTVGASDYHQHGSN TFLPQAQSGIYHVKGQPGQCPQPLPNLVETSCPNDMPKLCPADPWIAFHSHN DIWTSNVTREERLTVHNELANMEEDARSAGVATVLOEPDRYSGYWCPKAETI PSCKILLILAEENDSEVEI IHTVSPMLETRADSFYPKTGTANKPVTFKMSEIMI DAEGRIDIDKEILQIFELLFPGVETIARAGTPEGTAKWSILLDRSKRLRQLVLIS PELFIPVEDDVMERKILTESVSPVTLYIETTDIWINIHDIHFVPGSHEEIEF IFASECTGFRHLIKTILSKESKYRSSGGLPAPSDFKPIKEEIAITSGSEVILGR HGSIHQVDEVRLLVYFEGTKDSPLEHLVWYSVNPGBVTRLTDGQYSHSCCISQHC PFISKYNSKNPCHVCSLYKLSSPEDDPTCKTEFAWTLDSAGPLDPTPTPELFSFES TTGFTLYMLYKLDQPGKKYIFVLVFIYGGPQVLVNNRFGVKYIFRLMILASLGV VVVDNRGSHRGLEKGFAPKMGQVLEIDDQVGLQYLSRQYFIDLDVRVHGWSY GYDLSLMAWMORSDFRVAIAGAPVTLMIFYDTGYTERYMGHPDQNEQGYLGSVAMQ		

```
Db 967 AGCAATCTTAAGTCACTTTTAAAGATGTCAGAAATATATGATGATGCTGAAGAGGAT 1026
Qy 1233 CATAGATGTCATAGATGAAGAACTAATTCAACTTTTGAGATCTTATTTGAAGAGTTGA 1292
Db 1027 CATAGATGTCATAGATGAAGAACTAATTCAACTTTTGAGATCTTATTTGAAGAGTTGA 1086
Qy 1293 ATATATTGCCAGAGCTGGATGGACTCTCTGAGGAGAAATATGCTTGGTCCATCCCTACTAGA 1352
Db 1087 ATATATTGCCAGAGCTGGATGGACTCTCTGAGGAGAAATATGCTTGGTCCATCCCTACTAGA 1146
Qy 1353 TCGCTCCAGAGCTGCCCTACAGATAGTGTGATCTCACTGAAATATTTATTCACAGTAGA 1412
Db 1147 TCGCTCCAGAGCTGCCCTGAGATAGTGTGATCTCACTGAAATATTTATTCACAGTAGA 1206
Qy 1413 AGATGATGTTATGGAAGGAGAGACTCAATTGAGTCAGTGCCTGATCTGTGAGCCCACT 1472
Db 1207 AGATGATGTTATGGAAGGAGAGACTCAATTGAGTCAGTGCCTGATCTGTGAGCCCACT 1266
Qy 1473 AATTATCTATGAAGAAACAAACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTTT 1532
Db 1267 AATTATCTATGAAGAAACAAACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTTT 1326
Qy 1533 TCCCAAAAGTCACGAAGAGAAATTCAGTTTTATTTTGGCTCTGAAATGCAAAAACAGGTTT 1592
Db 1327 TCCCAAAAGTCACGAAGAGAAATTCAGTTTTATTTTGGCTCTGAAATGCAAAAACAGGTTT 1386
Qy 1593 CCGTCATTTATACAAAATACATCTATTTTAAAGGAAGCAAAATATAAACCAGTCCAGTGG 1652
Db 1387 CCGTCATTTATACAAAATACATCTATTTTAAAGGAAGCAAAATATAAACCAGTCCAGTGG 1446
Qy 1653 TGGCTCGCTGCTCCAGTGATTTCAAGTGTCTCTCAAGGAGGATAGCAATTCACAG 1712
Db 1447 TGGCTCGCTGCTCCAGTGATTTCAAGTGTCTCTCAAGGAGGATAGCAATTCACAG 1506
Qy 1713 TGGTGAATGGGAAGTCTTGGCCGCGATGGATCTAATATCCAAAGTGAAGTCAGAAG 1772
Db 1507 TGGTGAATGGGAAGTCTTGGCCGCGATGGATCTAATATCCAAAGTGAAGTCAGAAG 1566
Qy 1773 GCTGGTATTTTGAAGGCCAAGAACTCCCTTTAGAGCATACCTGTACGTAGTACAG 1832
Db 1567 GCTGGTATTTTGAAGGCCAAGAACTCCCTTTAGAGCATACCTGTACGTAGTACAG 1626
Qy 1833 TTACGTAAATCTGGAGAGTGACAAGGCTGACTGACCGTGGCTACTCACATTTCTGCTG 1892
Db 1627 TTACGTAAATCTGGAGAGTGACAAGGCTGACTGACCGTGGCTACTCACATTTCTGCTG 1886
Qy 1893 CATCAGTCAGACCTGTGACTCTTTTATAAGTATAGTAAACCAAGAAATCCACACTG 1952
Db 1687 CATCAGTCAGACCTGTGACTCTTTTATAAGTATAGTAAACCAAGAAATCCACACTG 1746
Qy 1953 TGTGTCCCTTTACAGCTATCAAGTCTCTGAAGTACCCCACTTGCAAAAACAAAGGAATT 2012
Db 1747 TGTGTCCCTTTACAGCTATCAAGTCTCTGAAGTACCCCACTTGCAAAAACAAAGGAATT 1806
Qy 2013 TTGGGCCACCATTTTGGATTCAGCAGGTCTCTTCTGACTATATCTCTCCAGAAATTTT 2072
Db 1807 TTGGGCCACCATTTTGGATTCAGCAGGTCTCTTCTGACTATATCTCTCCAGAAATTTT 1866
Qy 2073 CTCCTTTGAAAGTACTACTGATTTACATTTGATGGATGCTCTACAAGCCTCATGATCT 2132
Db 1867 CTCCTTTGAAAGTACTACTGATTTACATTTGATGGATGCTCTACAAGCCTCATGATCT 1926
Qy 2133 ACAGCTTGAAGAAATATCTACTGTCTGTTTCATATATATGTTGTTCTCAGGTGCAAGTT 2192
Db 1927 ACAGCTTGAAGAAATATCTACTGTCTGTTTCATATATATGTTGTTCTCAGGTGCAAGTT 1986
Qy 2193 GGTGAATAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTTAGCTCTCTAGG 2252
Db 1987 GGTGAATAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTTAGCTCTCTAGG 2046
Qy 2253 TTATGTGGTTGATGATGACAAACAGGGGATCTGTGTCACCGAGGCTTAAATTTGAAG 2312
```

```
Db 2047 TTATGTGGTTGATGATGACAAACAGGGGATCTGTCTACCGAGGGCTTAAATTTGAAG 2106
Qy 2313 CGCCTTTAAATATATAAATGGTCAAAATAGAAATGACGATCAGGTGGAGGACTCCCAATA 2372
Db 2107 CGCCTTTAAATATATAAATGGTCAAAATAGAAATGACGATCAGGTGGAGGACTCCCAATA 2166
Qy 2373 TCTAGCTTCTCGATATGATTTCAATTGACCTTAGATCGTGTGGGCATCCACGGCTGGTCTTA 2432
Db 2167 TCTAGCTTCTCGATATGATTTCAATTGACCTTAGATCGTGTGGGCATCCACGGCTGGTCTTA 2226
Qy 2433 TGGAGGATACCTCTCCCTGATGGCATTAAATGACAGAGGTGAGATATCTTTCAAGGTTGCTAT 2492
Db 2227 TGGAGGATACCTCTCCCTGATGGCATTAAATGACAGAGGTGAGATATCTTTCAAGGTTGCTAT 2286
Qy 2493 TGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAAACGTTATAT 2552
Db 2287 TGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAAACGTTATAT 2346
Qy 2553 GGGTCACCCCTGACACAGAAATGAAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGA 2612
Db 2347 GGGTCACCCCTGACACAGAAATGAAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGA 2406
Qy 2613 AAAGTTCCTCTGAAACCAAAATCGTTTACTGTCTTTACATGTTTCTCTGGATGAGAATGT 2672
Db 2407 AAAGTTCCTCTGAAACCAAAATCGTTTACTGTCTTTACATGTTTCTCTGGATGAGAATGT 2466
Qy 2673 CCATTTTGCACATACACAGTATATTACTTGAGTTTTTTTAGTGGGGCTGAAAAGCCATATGA 2732
Db 2467 CCATTTTGCACATACACAGTATATTACTTGAGTTTTTTTAGTGGGGCTGAAAAGCCATATGA 2526
Qy 2733 TTTACAGATCTATCTCCAGGAGACACAGCATAAGAGTTCTCTGAATCGGAGAACATTA 2792
Db 2527 TTTACAGATCTATCTCCAGGAGACACAGCATAAGAGTTCTCTGAATCGGAGAACATTA 2586
Qy 2793 TGAACCTGATCTTTTGCATCTACCTTACAGAAAACCTTGGATCAGTATTTGCTCTCTAAA 2852
Db 2587 TGAACCTGATCTTTTGCATCTACCTTACAGAAAACCTTGGATCAGTATTTGCTCTCTAAA 2646
Qy 2853 AGTGATAT 2860
Db 2647 AGTGATAT 2654

RESULT 11
AR631278 LOCUS AR631278 2671 bp DNA linear PAT 14-FEB-2005
DEFINITION Sequence 2 from patent US 6844180.
ACCESSION AR631278
VERSION AR631278.1 GI:59770920
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 2671)
AUTHORS Oi,S., Akinsanya,K.O., Riviere,P.J.M. and Junien,J.-L.
TITLE Serine protease genes related to DPPIV
JOURNAL Patent: US 6844180-A 2 18-JAN-2005;
Ferring BV;;
NLX;
FEATURES
source Location/Qualifiers
1..2671
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 80.0%; Score 2495; DB 6; Length 2671;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2645; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 213 CATGGCAGCAGCAATGGAACAGACAGCTGGGTGTGAGATATTTGAAACTGCGGACTG 272
Db 7 CATGGCAGCAGCAATGGAACAGACAGCTGGGTGTGAGATATTTGAAACTGCGGACTG 66
```

QY 273 TdAGGAGATATTGAATCACAGATCGGCCTAAATTTGGAGCCTTTTATGTTGAGCGGTA 332
Db |||||
QY 67 TGAGGAGATATTGAATCACAGATCGGCCTAAATTTGGAGCCTTTTATGTTGAGCGGTA 126
Db |||||
QY 333 TtCCTGGAGTCAGCTTAAAGCTGCTTGGCGATACAGAGAAATATCATGGCTACATGAT 392
Db |||||
QY 127 TtCCTGGAGTCAGCTTAAAGCTGCTTGGCGATACAGAGAAATATCATGGCTACATGAT 186
Db |||||
QY 393 GGTAGGACACCATGATTTTCATGTTTGAAGGGAATGATCCAGATGGACCTCATTC 452
Db |||||
QY 187 GGTAGGACACCATGATTTTCATGTTTGAAGGGAATGATCCAGATGGACCTCATTC 246
Db |||||
QY 453 AGACAGAACTTATTCCTTGGCCATGCTGGTGAAGACAGAGAAATACATGTTTATTC 512
Db |||||
QY 247 AGACAGAACTTATTCCTTGGCCATGCTGGTGAAGACAGAGAAATACATGTTTATTC 306
Db |||||
QY 513 TGAATTCCTTGAATTCATGATAGAGAGAGCTTAAATGCTCTCTTGGAGCCTCTTTT 572
Db |||||
QY 307 TGAATTCCTTGAATTCATGATAGAGAGAGCTTAAATGCTCTCTTGGAGCCTCTTTT 366
Db |||||
QY 573 GATCTTTTTCAGGCAACACTGACTATGGAATGATTTCTCGAAGAGAACTATTAAAG 632
Db |||||
QY 367 GATCTTTTTCAGGCAACACTGACTATGGAATGATTTCTCGAAGAGAACTATTAAAG 426
Db |||||
QY 633 AdAAGAGAAACGCAITGGAACTCGGAATTCCTTTACGATTTATCACCAGAGAGTGG 692
Db |||||
QY 427 AdAAGAGAAACGCAITGGAACTCGGAATTCCTTTACGATTTATCACCAGAGAGTGG 486
Db |||||
QY 693 AAcATTCTGTTTCAAGCGGTAGTGGAAATTTATCAGCTAAAGATGGAGGCGCACAGG 752
Db |||||
QY 487 AAcATTCTGTTTCAAGCGGTAGTGGAAATTTATCAGCTAAAGATGGAGGCGCACAGG 546
Db |||||
QY 753 ATTTCAGCAACACCTTTTAAGGCCCACTAGTGGAACTAGTTGTCCCAACATACGGAT 812
Db |||||
QY 547 ATTTCAGCAACACCTTTTAAGGCCCACTAGTGGAACTAGTTGTCCCAACATACGGAT 606
Db |||||
QY 813 GdATCCAAATATTATGCCCGCTGATCCAGACTGGATTCCTTTTATACATAGCAACGATAT 872
Db |||||
QY 607 GdATCCAAATATTATGCCCGCTGATCCAGACTGGATTCCTTTTATACATAGCAACGATAT 666
Db |||||
QY 873 TtGGATATCTAAACATCGTAAACAGAGAGAGAGACTCATTTATGTCACATGAGCT 932
Db |||||
QY 667 TtGGATATCTAAACATCGTAAACAGAGAGAGAGACTCATTTATGTCACATGAGCT 726
Db |||||
QY 933 AdCCAACTGGAAGAGATGCCAGATCAGCTGAGTCCCTACCTTTGTTCTCCAAAGA 992
Db |||||
QY 727 AdCCAACTGGAAGAGATGCCAGATCAGCTGAGTCCCTACCTTTGTTCTCCAAAGA 786
Db |||||
QY 993 AftTGATAGATATTCTGGCTATTGGTGTCCAAAGCTGAAACAACTCCAGTGGTG 1052
Db |||||
QY 787 AftTGATAGATATTCTGGCTATTGGTGTCCAAAGCTGAAACAACTCCAGTGGTG 846
Db |||||
QY 1053 TAAATTTCTTAGAATCTATACAGAAATGATGATGAGTGGGAAATTTATTCATGT 1112
Db |||||
QY 847 TAAATTTCTTAGAATCTATACAGAAATGATGATGAGTGGGAAATTTATTCATGT 906
Db |||||
QY 1113 TAcATCCCTATGTTTGAACAAGAGGCGAGATTCATTCGGTTATCTTAAACAGAGTAC 1172
Db |||||
QY 907 TAcATCCCTATGTTTGAACAAGAGGCGAGATTCATTCGGTTATCTTAAACAGAGTAC 966
Db |||||
QY 1173 AdCAAATCTTAAAGTCACTTTTAAAGATGTCAGAAATATGATGATGCTGAAGAGAGAT 1232
Db |||||
QY 967 AdCAAATCTTAAAGTCACTTTTAAAGATGTCAGAAATATGATGATGCTGAAGAGAGAT 1026
Db |||||
QY 1233 CMTAGATGTCATAGATAGAACTAATTCACCTTTTGAGATTCATTTGAGAGGTTGA 1292
Db |||||
QY 1027 CATAGATGTCATAGATAGAACTAATTCACCTTTTGAGATTCATTTGAGAGGTTGA 1086
Db |||||
QY 1293 ATATATTGCCAGAGCTGGATGGACTCCTGAGGAAATATGCTTGGTCCATCTACTAGA 1352
Db |||||
QY 1087 ATATATTGCCAGAGCTGGATGGACTCCTGAGGAAATATGCTTGGTCCATCTACTAGA 1146
Db |||||
QY 1353 TCGCTCCAGAGCTCGCCTACAGATAGTGTGATCTCACCTGAAATTTATTTATCCCAAGTAGA 1412
Db |||||

Db 1147 TCCTCCAGACTCCCTCGAGATAGTGTGATCTCACCTGAAATTTATTTATCCAGTAGA 1206
QY |||||
QY 1413 AGATGATGTTATTGGAAGGCAGAGACTCATTTGAGTCAGTGCCTGATTTCTGTGACGCAC 1472
Db |||||
QY 1207 AGATGATGTTATTGGAAGGCAGAGACTCATTTGAGTCAGTGCCTGATTTCTGTGACGCAC 1266
Db |||||
QY 1473 AATTATCTATGAAGAAACAAACAGACATCTGGATAAATATCCATGACATCTTTATGTTTT 1532
Db |||||
QY 1267 AATTATCTATGAAGAAACAAACAGACATCTGGATAAATATCCATGACATCTTTATGTTTT 1326
Db |||||
QY 1533 TCCCAAGCTCAGAGAGGAAATTTAGTTATTTTGGCCTCTCAATTCGAAACAGAGTTT 1592
Db |||||
QY 1327 TCCCAAGCTCAGAGAGGAAATTTAGTTATTTTGGCCTCTCAATTCGAAACAGAGTTT 1386
Db |||||
QY 1593 CCGTCATTTATACAAATTTACATCTATTTTAAAGSAAAGCAAAATATATAACGATCCAGTGG 1652
Db |||||
QY 1387 CCGTCATTTATACAAATTTACATCTATTTTAAAGSAAAGCAAAATATAAACGATCCAGTGG 1446
Db |||||
QY 1653 TGGGCTGCTGCTCCAAAGTGATTTCAAGTGTCTTCAAGAGGAGATAGCAATTACCAG 1712
Db |||||
QY 1447 TGGGCTGCTGCTCCAAAGTGATTTCAAGTGTCTTCAAGAGGAGATAGCAATTACCAG 1506
Db |||||
QY 1713 TGTGATGGAAGTTCCTTGGCGCATGATCTAATATCCAAGTTGATGAAGTCAGAAG 1772
Db |||||
QY 1507 TGTGATGGAAGTTCCTTGGCGCATGATCTAATATCCAAGTTGATGAAGTCAGAAG 1566
Db |||||
QY 1773 GCTGGTATATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACTGTACGTAGTCAG 1832
Db |||||
QY 1567 GCTGGTATATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACTGTACGTAGTCAG 1626
Db |||||
QY 1833 TTAGCTAAATCTCGAGAGAGTGACAAAGGTGACTGACCGTGGCTACTCACATTTCTGCTG 1892
Db |||||
QY 1627 TTAGCTAAATCTCGAGAGAGTGACAAAGGTGACTGACCGTGGCTACTCACATTTCTGCTG 1686
Db |||||
QY 1893 CATCAGTCAGCAGCTGAGCTTCTTTATAGTAAGTATAGTAACAGAGAAATCCACACTG 1952
Db |||||
QY 1687 CATCAGTCAGCAGCTGAGCTTCTTTATAGTAAGTATAGTAACAGAGAAATCCACACTG 1746
Db |||||
QY 1953 TGTGCTCCTTTTCAAGCTATCAAGTCTCTGAAGATGACCCAACTTGTCAAAAACAAAGGAATT 2012
Db |||||
QY 1747 TGTGCTCCTTTTCAAGCTATCAAGTCTCTGAAGATGACCCAACTTGTCAAAAACAAAGGAATT 1806
Db |||||
QY 2013 TTAGGCGCACATTTTGGATTCAGCAGGTCTCTTCTGACTATACTCTCCAGAAATTTT 2072
Db |||||
QY 1807 TTAGGCGCACATTTTGGATTCAGCAGGTCTCTTCTGACTATACTCTCCAGAAATTTT 1866
Db |||||
QY 2073 CTCTTTTGAAGTACTACTGGATTTTACATGTTATGGGATGCTCTACAAAGCCTCATGATCT 2132
Db |||||
QY 1867 CTCTTTTGAAGTACTACTGGATTTTACATGTTATGGGATGCTCTACAAAGCCTCATGATCT 1926
Db |||||
QY 2133 ACAGCCTGGAAGAAATATCCTACTCTGTGTTTCATATATGTTGTTCTCAGGTGCGAGTT 2192
Db |||||
QY 1927 ACAGCCTGGAAGAAATATCCTACTCTGTGTTTCATATATGTTGTTCTCAGGTGCGAGTT 1986
Db |||||
QY 2193 GGTGAATATTCGGTTTAAAGGAGTCAAGTAATTCGGCTTGAATACCTAGCCTCTCTAGG 2252
Db |||||
QY 1987 GGTGAATATTCGGTTTAAAGGAGTCAAGTAATTCGGCTTGAATACCTAGCCTCTCTAGG 2046
Db |||||
QY 2253 TTATGTGTTGTAGTAGATAGACAAAGGGGATCCTGTCAACGAGGCTTAAATTTGAAG 2312
Db |||||
QY 2047 TTATGTGTTGTAGTAGATAGACAAAGGGGATCCTGTCAACGAGGCTTAAATTTGAAG 2106
Db |||||
QY 2313 CGCCTTTAAATATAAAATGGGTCAAAATGATGATGATGAGTGGAGGAGTCCCAATA 2372
Db |||||
QY 2107 CGCCTTTAAATATAAAATGGGTCAAAATGATGATGATGAGTGGAGGAGTCCCAATA 2166
Db |||||
QY 2373 TCTAGCTTCGATATGATTTTCAATTCAGCTAGATGCTGGGATCCAGGCTGGTCTCTA 2432
Db |||||
QY 2167 TCTAGCTTCGATATGATTTTCAATTCAGCTAGATGCTGGGATCCAGGCTGGTCTCTA 2226
Db |||||
QY 2433 TGGAGGATACCTCTCCCTGATGTCATTAATGACAGAGTTCAGATATCTTTCAAGGTTGCTAT 2492
Db |||||

Db	2227	TGAGGATACCTCTCCCTGATGGCAATTAATGCAGAGGTGAGATATCTTTCAAGGTTGCTAT	2286	Db	187	GGCTAAGGACCACCATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGACCTCATTC	246
Qy	2493	TGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAACTGTATAT	2552	Qy	453	AGACAGAACTCTATTTACCTCCCATGCTCGTGAGAAACAGAGAAAATACACTGTTTTTATTC	512
Db	2287	TGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAACTGTATAT	2346	Db	247	AGACAGAACTCTATTTACCTCCCATGCTCGTGAGAAACAGAGAAAATACACTGTTTTTATTC	306
Qy	2553	GGGTCAACCTTGACACAGAAATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGA	2612	Qy	513	TGAAATTTCCCAAACCTATCAATAGACAGAGTCTTTAATGCTCTCTTGGAAGCCTCTTTT	572
Db	2347	GGGTCAACCTTGACACAGAAATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGA	2406	Db	307	TGAAATTTCCCAAACCTATCAATAGACAGAGTCTTTAATGCTCTCTTGGAAGCCTCTTTT	366
Qy	2613	AAAGTTCCTCTGACCAATCGTTTACTGCTCTTACATGCTGTTCTCGATGAGATGT	2672	Qy	573	GGATCTTTTTCAGGCAACACTGGACTATCGAATGTATTTCTCGAGAAGAGAACTATTAAAG	632
Db	2407	AAAGTTCCTCTGACCAATCGTTTACTGCTCTTACATGCTGTTCTCGATGAGATGT	2466	Db	367	GGATCTTTTTCAGGCAACACTGGACTATCGAATGTATTTCTCGAGAAGAGAACTATTAAAG	426
Qy	2673	CCATTTTGCATACACAGATATTTACTGATTTTTTTAGTAGGGCTGAAAGCCATATGA	2732	Qy	633	AGAAAGAAAAACGCAITGGAAACAGTCGGAATTCCTTTACGATTTATCACCGTAAAGAGATGG	692
Db	2467	CCATTTTGCATACACAGATATTTACTGATTTTTTTAGTAGGGCTGAAAGCCATATGA	2526	Db	427	AGAAAGAAAAACGCAITGGAAACAGTCGGAATTCCTTTACGATTTATCACCGTAAAGAGATGG	486
Qy	2733	TTTACAGATCTATCTCCAGAGAGACACAGCATTAAGATTTCTGAATCGGGAGAACATTA	2792	Qy	693	AACATTTCTGTTTCAAGCCGGTAGTGGAAATTTATCACGTAAGAGATGGAGGCCACAAAGG	752
Db	2527	TTTACAGATCTATCTCCAGAGAGACACAGCATTAAGATTTCTGAATCGGGAGAACATTA	2586	Db	487	AACATTTCTGTTTCAAGCCGGTAGTGGAAATTTATCACGTAAGAGATGGAGGCCACAAAGG	546
Qy	2793	TGAATCGATCTTTTGACCTACCTTCAAGAAAACCTTGGATCACGTATTCGCTCTTAA	2852	Qy	753	ATTTACGCAACAACTTTTAAGGCCAACTAGTGGAAACTAGTTGTCCCAACATACCGAT	812
Db	2587	TGAATCGATCTTTTGACCTACCTTCAAGAAAACCTTGGATCACGTATTCGCTCTTAA	2646	Db	547	ATTTACGCAACAACTTTTAAGGCCAACTAGTGGAAACTAGTTGTCCCAACATACCGAT	606
Qy	2853	AGTGATAT 2860		Qy	813	GGATCCAAAATTTATGCCCTGCTGATCCAGACTGGATTCGTTTATACATAGCAACGATAT	872
Db	2647	AGTGATAT 2654		Db	607	GGATCCAAAATTTATGCCCTGCTGATCCAGACTGGATTCGTTTATACATAGCAACGATAT	666
RESULT 12				Qy	873	TTGGATATCTAACATCGTAACCCAGAGAGAGAGAGACTCACTTATGTCACATAGAGCT	932
AX608725				Db	667	TTGGATATCTAACATCGTAACCCAGAGAGAGAGAGACTCACTTATGTCACATAGAGCT	726
LOCUS	AX608725	Sequence 2 from Patent WO0231134.	linear	Qy	933	AGCCAAACATGGAGAGAGATCCAGATCAGCTGGAGTGGCTACCTTGTTCCTCCAGAGAA	992
DEFINITION	AX608725	AX608725.1 GI:28404298		Db	727	AGCCAAACATGGAGAGAGATCCAGATCAGCTGGAGTGGCTACCTTGTTCCTCCAGAGAA	786
ACCESSION				Qy	993	ATTTGATAGATATTTGGCTATTTGGTGTCCTCCAAAGCTGAAACAACCTCCAGTGGTG	1052
VERSION				Db	787	ATTTGATAGATATTTGGCTATTTGGTGTCCTCCAAAGCTGAAACAACCTCCAGTGGTG	846
KEYWORDS				Qy	1053	TAAATTTCTTAGAATCTATATGAAGAAAATGATGAATCTGAGGTGGAATTTATTCATGT	1112
SOURCE				Db	847	TAAATTTCTTAGAATCTATATGAAGAAAATGATGAATCTGAGGTGGAATTTATTCATGT	906
ORGANISM				Qy	1113	TACATCCCTATGTTGGAAAACAGAGGGCAGATTCATTCGTTTATCCTAAAAACAGGTAC	1172
REFERENCE				Db	907	TACATCCCTATGTTGGAAAACAGAGGGCAGATTCATTCGTTTATCCTAAAAACAGGTAC	966
AUTHORS				Qy	1173	AGCAAAATCCTAAAGTCACCTTTTAAAGATGTCAGAAAATTAATGATTGATGCTGGAAGGAAGGAT	1232
TITLE				Db	967	AGCAAAATCCTAAAGTCACCTTTTAAAGATGTCAGAAAATTAATGATTGATGCTGGAAGGAAGGAT	1026
JOURNAL				Qy	1233	CATAGATGTCATAGATAAGGAACCTAAATTCACCTTTTGGAGATTTCTATTTCAAGGAGTTGA	1292
FEATURES				Db	1027	CATAGATGTCATAGATAAGGAACCTAAATTCACCTTTTGGAGATTTCTATTTCAAGGAGTTGA	1086
source				Qy	1293	ATATATTTGCAGAGCTGGATGGACTCTCGAGGAAAAATATGCTTGGTCCATCTACTAGA	1352
1				Db	1087	ATATATTTGCAGAGCTGGATGGACTCTCGAGGAAAAATATGCTTGGTCCATCTACTAGA	1146
Q1,S., Akinaanya, K.O., Riviere, P.J. and Junien, J.L.				Qy	1353	TGCTCCCAAGACTCCGCTTACAGATAGTGTGATCTCACCTGAAATTTATTTATCCAGTGA	1412
Novel serine protease genes related to dppiv				Db	1147	TGCTCCCAAGACTCCGCTTACAGATAGTGTGATCTCACCTGAAATTTATTTATCCAGTGA	1206
Patent: WO 0231134-A 2 18-APR-2002;				Qy	1413	AGATGATGTTATGGAAGGCGAGAGACTCATTTGAGTCACTGCTGATTTCTGTGACGCACT	1472
Ferring BV (NL)				Db	1207	AGATGATGTTATGGAAGGCGAGAGACTCATTTGAGTCACTGCTGATTTCTGTGACGCACT	1266
Location/Qualifiers				Qy	1473	AATTTATCTATGAAGAAAACACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTT	1532
1				Db	1267	AATTTATCTATGAAGAAAACACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTT	1326
/organism="Homo sapiens"							
/mol_type="unassigned DNA"							
/db_xref="taxon:9606"							
Query Match							
Best Local Similarity							
Matches 2645; Conservative							
0; Mismatches							
3; Indels							
0; Gaps							
0;							
Qy	213	CATGGCAGCAGCAATGGAAAACAGAACAGCTGGGTGTTGAGATATTTGAAACTCCGACTG	272				
Db	7	CATGGCAGCAGCAATGGAAAACAGAACAGCTGGGTGTTGAGATATTTGAAACTCCGACTG	66				
Qy	273	TGAGGAGAAATTTGAATCACAGGATCGGCCCTAAATTTGGAGCCCTTTTATGTTGAGCGGTA	332				
Db	67	TGAGGAGAAATTTGAATCACAGGATCGGCCCTAAATTTGGAGCCCTTTTATGTTGAGCGGTA	126				
Qy	333	TTCTGGAGTCACTTAAAAAGCTGCTTGGCGATACAGAAAATATATCATGGCTACATGAT	392				
Db	127	TTCTGGAGTCACTTAAAAAGCTGCTTGGCGATACAGAAAATATATCATGGCTACATGAT	186				
Qy	393	GGCTAAGGCACCATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGGACCTCATTC	452				

Qy	1533	TCGCCAAGTCAAGAGGAAATTGAGTTTATTTTGGCTCTGAATGCAAAACAGGTTT	1592
Db	1327	TCGCCAAGTCAAGAGGAAATTGAGTTTATTTTGGCTCTGAATGCAAAACAGGTTT	1386
Qy	1593	CGGTCAATTTATCAAAATTTACATCTATTTTAAAGGAAAGCAAAATATAAAGCGATCCAGTGG	1652
Db	1387	CGGTCAATTTATCAAAATTTACATCTATTTTAAAGGAAAGCAAAATATAAAGCGATCCAGTGG	1446
Qy	1653	TGGGTGCTGCTGCCAAGTGATTTTCAAGTGTCCTTCAAAAGAGGAGATAGCAATTTACCAG	1712
Db	1447	TGGGTGCTGCTGCCAAGTGATTTTCAAGTGTCCTTCAAAAGAGGAGATAGCAATTTACCAG	1506
Qy	1713	TGGTGAATGGGAAGTCTTGGCCGCGCATGGAATCTAATATCAAGTTGATGAAGTCAGAG	1772
Db	1507	TGGTGAATGGGAAGTCTTGGCCGCGCATGGAATCTAATATCAAGTTGATGAAGTCAGAG	1566
Qy	1773	GGTGGTATATTTTGAAGGACCAAAAGACTCCCTTTAGAGCATCACTGTACGTAGTCAG	1832
Db	1567	GGTGGTATATTTTGAAGGACCAAAAGACTCCCTTTAGAGCATCACTGTACGTAGTCAG	1626
Qy	1833	TTACGTAAATCTCGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCAATTTTGCTG	1892
Db	1627	TTACGTAAATCTCGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCAATTTTGCTG	1686
Qy	1893	CATCAGTCAGCACTGTGACTCTTTATTAAGTAAGTATAGTAACCCAGAGAATCCACACTG	1952
Db	1687	CATCAGTCAGCACTGTGACTCTTTATTAAGTAAGTATAGTAACCCAGAGAATCCACACTG	1746
Qy	1953	TGTGTCCTTTTCAAGCTATCAAGTCTCTGAAGATGACCCAACTTGCAGAAACAAAGAAAT	2012
Db	1747	TGTGTCCTTTTCAAGCTATCAAGTCTCTGAAGATGACCCAACTTGCAGAAACAAAGAAAT	1806
Qy	2013	TTGGGCACATTTTGGATTCAGAGGCTCTCTCTCGACTATGACTCTCTCAGAAATTTT	2072
Db	1807	TTGGGCACATTTTGGATTCAGAGGCTCTCTCTCGACTATGACTCTCTCAGAAATTTT	1866
Qy	2073	CTCTTTTGAAGTACTGATGATTTTACATTTGATGGATGCTTACAGGCTCATGATCT	2132
Db	1867	CTCTTTTGAAGTACTGATGATTTTACATTTGATGGATGCTTACAGGCTCATGATCT	1926
Qy	2133	ACAGCTGGAAAGAAATATCTTACTGTCTGTCTATATATGTTGGTCTCTCAGGTGCAAT	2192
Db	1927	ACAGCTGGAAAGAAATATCTTACTGTCTGTCTATATATGTTGGTCTCTCAGGTGCAAT	1986
Qy	2193	GGTGAATATTCGGTTTAAAGAGTCAAGTATTTTCGGCTTGAATACCTTAGCCTCTTAGG	2252
Db	1987	GGTGAATATTCGGTTTAAAGAGTCAAGTATTTTCGGCTTGAATACCTTAGCCTCTTAGG	2046
Qy	2253	TTATGTGTTGTAGTATAGACAACAGGGGATCTGTCTCCAGAGGCTTAAATTTGAAGG	2312
Db	2047	TTATGTGTTGTAGTATAGACAACAGGGGATCTGTCTCCAGAGGCTTAAATTTGAAGG	2106
Qy	2313	CGCTTTAAATATAAAATGGGTCAATAGAAATTCAGCATCAGGTGGAGAGCTCCAAATA	2372
Db	2107	CGCTTTAAATATAAAATGGGTCAATAGAAATTCAGCATCAGGTGGAGAGCTCCAAATA	2166
Qy	2373	TCTAGCTTCTCGATATGATTTCAATGACTTTAGATCGTGGGCATCCACGGCTGGTCTTA	2432
Db	2167	TCTAGCTTCTCGATATGATTTCAATGACTTTAGATCGTGGGCATCCACGGCTGGTCTTA	2226
Qy	2433	TGGAGGATACCTCTCCCTGATGGCATTAAATGCAGAGGTGAGATCTTTCAGGGTTGCTAT	2492
Db	2227	TGGAGGATACCTCTCCCTGATGGCATTAAATGCAGAGGTGAGATCTTTCAGGGTTGCTAT	2286
Qy	2493	TGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACAGGAACCTTATAT	2552
Db	2287	TGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACAGGAACCTTATAT	2346
Qy	2553	GGGTCAACCTGACAGAAATGAACAGGGCTATTTACTTAGGATCTGTGGCCATGCAAGCAGA	2612
Db	2347	GGGTCAACCTGACAGAAATGAACAGGGCTATTTACTTAGGATCTGTGGCCATGCAAGCAGA	2406

Qy	2613	AAAGTTCCCTCTGAACCAAAATCGTTTACTGCTCTTACATGGTTTCTTGGATGAGAATGT	2672
Db	2407	AAAGTTCCCTCTGAACCAAAATCGTTTACTGCTCTTACATGGTTTCTTGGATGAGAATGT	2466
Qy	2673	CCATTTTGCACATACCAAGTATATTACTGAGTTTTTTTGTAGTGGGCTGGAAAGCCATATGA	2732
Db	2467	CCATTTTGCACATACCAAGTATATTACTGAGTTTTTTTGTAGTGGGCTGGAAAGCCATATGA	2526
Qy	2733	TTTACAGATCTATCTCTCAGGAGAGACACAGCATAAAGAGTTTCTGAAATCGGAGAACATTA	2792
Db	2527	TTTACAGATCTATCTCTCAGGAGAGACACAGCATAAAGAGTTTCTGAAATCGGAGAACATTA	2586
Qy	2793	TGAACGTGACTCTTTTGCACCTACCTTCAAGAAAACCTTGGATCAGTATTCCTGCTCTAAA	2852
Db	2587	TGAACGTGACTCTTTTGCACCTACCTTCAAGAAAACCTTGGATCAGTATTCCTGCTCTAAA	2646
Qy	2853	AGTGATAT 2860	
Db	2647	AGTGATAT 2654	
RESULT 13			
AR448400			
LOCUS	AR448400	2797 bp	DNA linear PAT 20-FEB-2004
DEFINITION	Sequence 1103 from patent US 6,673,549.		
ACCESSION	AR448400		
VERSION	AR448400.1 GI:42676724		
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unclassified.		
REFERENCE	1 (bases 1 to 2797)		
AUTHORS	Furness, L.M. and Buchbinder, J.L.		
TITLE	Genes expressed in C3A liver cell cultures treated with steroids		
JOURNAL	Patent: US 6,673,549-A 1103 06-JAN-2004;		
	Incyte Corporation; Palo Alto, CA		
FEATURES	Location/Qualifiers		
source	1..2797		
	/organism="unknown"		
	/mol_type="genomic DNA"		
ORIGIN			
Query Match	74.2%;	Score 2315;	DB 6; Length 2797;
Best Local Similarity	99.8%;	Pred. No: 0;	
Matches 2785;	Conservative 0;	Mismatches 1;	Indels 5; Gaps 2;
Qy	335	CCTGGAGTCAGCTTAAAAAGCTGCTTGCAGATACCAAGAAATATCATGCTACATGATGG	394
Db	1	CCTGGAGTCAGCTTAAAAAGCTGCTTGCAGATACCAAGAAATATCATGCTACATGATGG	60
Qy	395	CTAAGGACCAACATGATTTTCATGTTGTGAAGAGGAATGATCCAGATGCACCTCATTCAG	454
Db	61	CTAAGGACCAACATGATTTTCATGTTGTGAAGAGGAATGATCCAGATGCACCTCATTCAG	120
Qy	455	ACAGATCTATTACCTTGCCTGCTGTGTGAGAACACAGAGAAAATACACTGTTTTATTCTG	514
Db	121	ACAGATCTATTACCTTGCCTGCTGTGTGAGAACACAGAGAAAATACACTGTTTTATTCTG	180
Qy	515	AAATTCCTCAAACTATCAATAGACAGCAGTCTTAATGCTCTCTTGGAAAGCCTCTTTTGG	574
Db	181	AAATTCCTCAAACTATCAATAGACAGCAGTCTTAATGCTCTCTTGGAAAGCCTCTTTTGG	240
Qy	575	ATCTTTTTCAGGCAACATCGGACTATGGAATGTGTTTCTCGAGAGAGAACTATTAAAG	634
Db	241	ATCTTTTTCAGGCAACATCGGACTATGGAATGTGTTTCTCGAGAGAGAACTATTAAAG	300
Qy	635	AAAGAAAACGCAATTGGAAACAGTCGGAATTCCTTCTTACGATTTATCACCAGGAAGTGGAA	694
Db	301	AAAGAAAACGCAATTGGAAACAGTCGGAATTCCTTCTTACGATTTATCACCAGGAAGTGGAA	360
Qy	695	CATTTCTGTTTCAAGCCGGTAGTGGAAATTTATCAGTAAAAAGATGGAGGGCCACAAGGAT	754
Db	361	CATTTCTGTTTCAAGCCGGTAGTGGAAATTTATCAGTAAAAAGATGGAGGGCCACAAGGAT	420

QY 755 TTAGCGCAACAACTTTAAGCCCAATCTAGTGGAACTAGTTGTGCCAAACATACGGATGG 814
Db 421 TTAGCGCAACAACTTTAAGCCCAATCTAGTGGAACTAGTTGTGCCAAACATACGGATGG 480
QY 815 ATCCAAAATATATGCCCGCTGTATCCAGACTGGATTGCTTTTATACATAGCAACGATATTT 874
Db 481 ATCCAAAATATATGCCCGCTGTATCCAGACTGGATTGCTTTTATACATAGCAACGATATTT 540
QY 875 GGATATCTAACTCGTAACCCAGAGAAAGAGAGACTCACTTATGTGCACAAATGAGCTAG 934
Db 541 GGATATCTAACTCGTAACCCAGAGAAAGAGAGACTCACTTATGTGCACAAATGAGCTAG 600
QY 935 CCAACATGGAAGAGATGCCAGATCAGCTGGATCGCTACCTTTGTTCTCCAGAGAAAT 994
Db 601 CCAACATGGAAGAGATGCCAGATCAGCTGGATCGCTACCTTTGTTCTCCAGAGAAAT 660
QY 995 TTGATAGATATCTGGCTATTCGGTGTCCAAAGCTGAAACAACTCCAGTGGTGGTA 1054
Db 661 TTGATAGATATCTGGCTATTCGGTGTCCAAAGCTGAAACAACTCCAGTGGTGGTA 720
QY 1055 AAATTTCTTGAATTTCTATATGAAGAAATGATGAATCTGAGGTGGAAATTAATTCATGTTA 1114
Db 721 AAATTTCTTGAATTTCTATATGAAGAAATGATGAATCTGAGGTGGAAATTAATTCATGTTA 780
QY 1115 CATCCCTATGTTGGAACAAGG-AGGCGAGATTCATTCGGTTATCCTAAAAACAGGTACA 1173
Db 781 CATCCCTATGTTGGAACAAGGCGAGGCGAGATTCATTCGGTTATCCTAAAAACAGGTACA 840
QY 1174 GCAATTCCTAAAGTCACTTTTAAAGATGTCAGAAATAATGATTCGTCGTAAGGAAGGATC 1233
Db 841 GCAATTCCTAAAGTCACTTTTAAAGATGTCAGAAATAATGATTCGTCGTAAGGAAGGATC 900
QY 1234 ATAGATGTCATAGATAAGGAACATAATCAACCTTTTGGAGATTTCTATTTGAAGGAGTTGAA 1293
Db 901 ATAGATGTCATAGATAAGGAACATAATCAACCTTTTGGAGATTTCTATTTGAAGGAGTTGAA 960
QY 1294 TATATTTGCCAGAGCTGGATGGACTCCTGAGGGAATAAGTGTGCTTGGTCCATCCTACT 1349
Db 961 TATATTTGCCAGAGCTGGATGGACTCCTGAGGGAATAAGTGTGCTTGGTCCATCCTACT 1020
QY 1350 AGATCGCTCCAGACTCCGCTACAGATAGTGTTCATCTCACCTGAAATTTATTTATCCCACT 1409
Db 1021 AGATCGCTCCAGACTCCGCTACAGATAGTGTTCATCTCACCTGAAATTTATTTCCCACT 1080
QY 1410 AGAAGATGATGTTATGGAAGGCGAGAGACTCATTTGAGTCAGTCCCTGATTCGTGACGCC 1469
Db 1081 AGAAGATGATGTTATGGAAGGCGAGAGACTCATTTGAGTCAGTCCCTGATTCGTGACGCC 1140
QY 1470 ACTTAATCTATGGAAGAAACACAGACATCTGGATTAATATCCATGACATCTTTTCATGT 1529
Db 1141 ACTTAATCTATGGAAGAAACACAGACATCTGGATTAATATCCATGACATCTTTTCATGT 1280
QY 1530 TTTTCCCAAGTCACGAAGAGGAATTTGAGTTATTTTGGCTCTGAAATGCAAAACAGG 1589
Db 1201 TTTTCCCAAGTCACGAAGAGGAATTTGAGTTATTTTGGCTCTGAAATGCAAAACAGG 1260
QY 1590 TTTCCGTCATTTATCAAAATTAATCTATTTTAAAGGAAGCAAAATATAACGATCCAG 1649
Db 1261 TTTCCGTCATTTATCAAAATTAATCTATTTTAAAGGAAGCAAAATATAACGATCCAG 1320
QY 1650 TGGTGGGCTGCTCTCCAGTGATTTCAAGTGTCTCTATCAAGAGGAGATAGCAATTAAC 1709
Db 1321 TGGTGGGCTGCTCTCCAGTGATTTCAAGTGTCTCTATCAAGAGGAGATAGCAATTAAC 1380
QY 1710 CAGTGGTGAATGGGAAGTTCTTGGCCGGCATGGATCTAATATCCAAAGTTGATGAAGTCAG 1769
Db 1381 CAGTGGTGAATGGGAAGTTCTTGGCCGGCATGGATCTAATATCCAAAGTTGATGAAGTCAG 1440
QY 1770 AAGGCTGGTATATTTTGAAGGCAACAAAGACTCCCTTTAGAGCATCACCTGTAGGTAGT 1829
Db 1441 AAGGCTGGTATATTTTGAAGGCAACAAAGACTCCCTTTAGAGCATCACCTGTAGGTAGT 1500

QY 1830 CAGTTACGTAAATCCCTGGAGGAGTGACAAGGCTGACTGACCGTCTACTACATTTCTTG 1889
Db 1501 CAGTTACGTAAATCCCTGGAGAGGTGACAAGGCTGACTGACCGTCTACTACATTTCTTG 1560
QY 1890 CTGCATCAGTCAGCACTGTGACTTCTTTATAAGTAAGTATAGTAAACGAGAAGATCCACA 1949
Db 1561 CTGCATCAGTCAGCACTGTGACTTCTTTATAAGTAAGTATAGTAAACGAGAAGATCCACA 1620
QY 1950 CTGTGTCTCCCTTTACAGCTATCAAGTCTCTGAAGATGACCCAACTTGCAAAAACAAAGGA 2009
Db 1621 CTGTGTCTCCCTTTACAGCTATCAAGTCTCTGAAGATGACCCAACTTGCAAAAACAAAGGA 1680
QY 2010 ATTTGGGCCACCACTTTTGGATTACAGAGTCTCTCTGACTATATCTCTCAGAGAAAT 2069
Db 1681 ATTTGGGCCACCACTTTTGGATTACAGAGTCTCTCTGACTATATCTCTCAGAGAAAT 1740
QY 2070 TTTCTCTTTTGAAGTACTACTGGATTTTACATTTGATGGGATGCTCTACAAGCCTCATGA 2129
Db 1741 TTTCTCTTTTGAAGTACTACTGGATTTTACATTTGATGGGATGCTCTACAAGCCTCATGA 1800
QY 2130 TCTACAGCCTGGAAAGAAATATCTACTGTGTCTCATATATATGTTGTTGCTCTCAGGTGCA 2189
Db 1801 TCTACAGCCTGGAAAGAAATATCTACTGTGTCTCATATATGTTGTTGCTCTCAGGTGCA 1860
QY 2190 GTTGGTGAATAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTTAGCCTCTCT 2249
Db 1861 GTTGGTGAATAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTTAGCCTCTCT 1920
QY 2250 AGGTTATGTGGTTGTAGTAGACACAACAGGGGATCTGTCCACCGAGGGCTTAAATTTGA 2309
Db 1921 AGGTTATGTGGTTGTAGTAGACACAACAGGGGATCTGTCCACCGAGGGCTTAAATTTGA 1980
QY 2310 AGCGCGCTTTAAATATAAATGAGTCAAAATAGAAATTTGACGATCAGTGGAGACTTCCA 2369
Db 1981 AGCGCGCTTTAAATATAAATGAGTCAAAATAGAAATTTGACGATCAGTGGAGACTTCCA 2040
QY 2370 ATATCTAGCTTCTCGATATGATTTTCATTTGACTTAGATCGTGTGGGCATCCACGGCTGGTC 2429
Db 2041 ATATCTAGCTTCTCGATATGATTTTCATTTGACTTAGATCGTGTGGGCATCCACGGCTGGTC 2100
QY 2430 CTATGAGAGATACCTCTCCCTGATGGCAATTAATCGAGAGTCCAGATATCTTTCAGGGTTGC 2489
Db 2101 CTATGAGAGATACCTCTCCCTGATGGCAATTAATCGAGAGTCCAGATATCTTTCAGGGTTGC 2160
QY 2490 TATTTGCTGGGCCCCCAGTCACTCTGTGGAICTTCTATGATACAGGATACACGGAAACGTTA 2549
Db 2161 TATTTGCTGGGCCCCCAGTCACTCTGTGGAICTTCTATGATACAGGATACACGGAAACGTTA 2220
QY 2550 TATGGGTCAACCTGACCAAGAAATGAACAGGGCTATTTACTTAGGATCTGTGGCCATGCAAGC 2609
Db 2221 TATGGGTCAACCTGACCAAGAAATGAACAGGGCTATTTACTTAGGATCTGTGGCCATGCAAGC 2280
QY 2610 AGAAAAGTTTCCCTCTGAAACCAAAATCGTTTACTGCTCTTACATGGTTTCTTGGATCAGAA 2669
Db 2281 AGAAAAGTTTCCCTCTGAAACCAAAATCGTTTACTGCTCTTACATGGTTTCTTGGATCAGAA 2340
QY 2670 TGTCCATTTTGCACATACAGTATATTTACTGAGTTTTTTTAGTGAGGCTGGAAAGCCATA 2729
Db 2341 TGTCCATTTTGCACATACAGTATATTTACTGAGTTTTTTTAGTGAGGCTGGAAAGCCATA 2400
QY 2730 TGAATTTACAGATCTATCTCAGGAGAGACACACATAGAGTTTCCCTGAAATCCGGAGAAC 2789
Db 2401 TGAATTTACAGATCTATCTCAGGAGAGACACACATAGAGTTTCCCTGAAATCCGGAGAAC 2460
QY 2790 TTATCAACTGATCTTTTGGCACTACCTTCAAGAAAACCTTGGATACAGTATTTGCTGCTCT 2849
Db 2461 TTATGAACTGATCTTTTGGCACTACCTTCAAGAAAACCTTGGATACAGTATTTGCTGCTCT 2520
QY 2850 AAAAGTGATATAATTTTGAACCTGTGTAGAACTCTCTGGGTATACACTGGGCTATTTAACCAA 2909
Db 2521 AAAAGTGATATAATTTTGAACCTGTGTAGAACTCTCTGGGTATACACTGGGCTATTTAACCAA 2580
QY 2910 ATGAGGAGTTTAAATCAACAGAAAAACAGAAATTTGATCATCATTTTGTATACCTGCCAT 2969

Db 2581 ATGAGGAGTTTAAATCAACAGAAAACACAGAAATTGATCATCACATTTTGATACCTGCCAT 2640
Qy 2970 GTAACATCTACTCTCGAAATAAATCTGGTGCATCGAGGGTCTACGGTTTGTGGTAGT 3029
Db 2641 GTAACATCTACTCTCGAAATAAATCTGGTGCATCGAGGGTCTACGGTTTGTGGTAGT 2700
Qy 3030 AATCTAATACCTTAACCCACATGCTCAAAATCAAATGATACATATTCCTGAGAGCCCA 3089
Db 2701 AATCTAATACCTTAACCCACATGCTCAAAATCAAATGATACATATTCCTGAGAGCCCA 2760
Qy 3090 GCAATACATAAGAAATTAATAAAAAAAA 3120
Db 2761 GCAATACATAAGAAATTAATAAAAAAAA 2791

RESULT 14
AR631281 4523 bp DNA linear PAT 14-FEB-2005
LOCUS Sequence 8 from patent US 6844180.
DEFINITION AR631281
ACCESSION AR631281
VERSION AR631281.1 GI:59770924
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
AUTHORS 1 (bases 1 to 4523)
TITLE Oi.S., Akinsanya, K.O., Riviere, P.J.M. and Junien, J.-L.
JOURNAL Serine protease genes related to DpPIV
Patent: US 6844180-A 8 18-JAN-2005;
Ferring BV;;
NLX; Location/Qualifiers
1..4523
/organism="unknown"
/mol_type="genomic DNA"

FEATURES
source

ORIGIN
Query Match 68.2%; Score 2128; DB 6; Length 4523;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2178; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AAGTGCTAAAGCCTCGAGGCCAAGCCGCTGCTACTGCCGCCGCTGCTTCTTAGTGCCG 60
Db 1 AAGTGCTAAAGCCTCGAGGCCAAGCCGCTGCTACTGCCGCCGCTGCTTCTTAGTGCCG 60
Qy 61 CGTTCGCCGCTGGGTTGTACCGCGCGCGCGCGAGGAGCCACTGCAACAGACCG 120
Db 61 CGTTCGCCGCTGGGTTGTACCGCGCGCGCGCGAGGAGCCACTGCAACAGACCG 120
Qy 121 GAGTGGAGGCGCGCAGCATGAAGCGCGCGCGCGCTCCATAGCCACGTGGGACCG 180
Db 121 GAGTGGAGGCGCGCAGCATGAAGCGCGCGCGCGCTCCATAGCCACGTGGGACCG 180
Qy 181 TCGGCGCGCGCGCGGGAAGAAATGCAACATGGCAGCAGCAATGGAAAACAGAACAG 240
Db 181 TCGGCGCGCGCGCGGGAAGAAATGCAACATGGCAGCAGCAATGGAAAACAGAACAG 240
Qy 241 CTGGGTGTCAGATATTTGAAACTGCGGACTGTGAGGAGAAATTTGAATCAGAGATCGG 300
Db 241 CTGGGTGTCAGATATTTGAAACTGCGGACTGTGAGGAGAAATTTGAATCAGAGATCGG 300
Qy 301 CCTAAATGGAGCGCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAGAGCTGTT 360
Db 301 CCTAAATGGAGCGCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAGAGCTGTT 360
Qy 361 GCCGATACAGAAAATATCATGCTACATGATGGCTTAAGGACACCATGATTTTCATGTTT 420
Db 361 GCCGATACAGAAAATATCATGCTACATGATGGCTTAAGGACACCATGATTTTCATGTTT 420
Qy 421 GTGAGAGGAAATGATCCAGATGGACCTCAATTCAGACAGAAATCTATTAACCTTGCAATGCT 480
Db 421 GTGAGAGGAAATGATCCAGATGGACCTCAATTCAGACAGAAATCTATTAACCTTGCAATGCT 480

Qy 481 GGTGAGAAACAGAGAAATACACTGTTTATCTGAAATTTCCAAAACTATCAATAGAGCA 540
Db 481 GGTGAGAAACAGAGAAATACACTGTTTATCTGAAATTTCCAAAACTATCAATAGAGCA 540
Qy 541 GCACTCTTAATGCTCTCTTTGGAAAGCCCTCTTTTGGATCTTTTTCAGGCAACACTGGACTAT 600
Db 541 GCACTCTTAATGCTCTCTTTGGAAAGCCCTCTTTTGGATCTTTTTCAGGCAACACTGGACTAT 600
Qy 601 GGAATGTTATCTCGAGAAAGAACTATTAAAGAGAAAGAAACGCAATTTGGAACAGTCGGA 660
Db 601 GGAATGTTATCTCGAGAAAGAACTATTAAAGAGAAAGAAACGCAATTTGGAACAGTCGGA 660
Qy 661 ATTGCTCTTACGATTTATCCAAAGAAAGTGGAAACATTTCTGTTTCAAGCCGGTAGTGA 720
Db 661 ATTGCTCTTACGATTTATCCAAAGAAAGTGGAAACATTTCTGTTTCAAGCCGGTAGTGA 720
Qy 721 ATTTATACGTTAAAGATGAGGGCCACAGGATTTTACGCAACAACTTTAAGGCCCAAT 780
Db 721 ATTTATACGTTAAAGATGAGGGCCACAGGATTTTACGCAACAACTTTAAGGCCCAAT 780
Qy 781 CTAGTGGAACTAGTGTCTCCCAACATACGATCGATCCAAATTTATGCCCCCTGATCCA 840
Db 781 CTAGTGGAACTAGTGTCTCCCAACATACGATCGATCCAAATTTATGCCCCCTGATCCA 840
Qy 841 GACTGGATTGCTTTTATACATAGCAACGATTTTGGATATCTTAACATCTGAACAGAGAA 900
Db 841 GACTGGATTGCTTTTATACATAGCAACGATTTTGGATATCTTAACATCTGAACAGAGAA 900
Qy 901 GAAAGAGACTCACTTATGTCACAAATGAGTAGCAACATGCAAGAAAGATGCCAGATCA 960
Db 901 GAAAGAGACTCACTTATGTCACAAATGAGTAGCAACATGCAAGAAAGATGCCAGATCA 960
Qy 961 GCTGGAGTCGCTACCTTTGTTCTCCAAAGAAATTTGATAGATATTTCTGGCTATTGGTGG 1020
Db 961 GCTGGAGTCGCTACCTTTGTTCTCCAAAGAAATTTGATAGATATTTCTGGCTATTGGTGG 1020
Qy 1021 TGTCCAAAAGCTGAAACAACTCCCAAGTGGTGGTAAAAATTTTGAATTTCTATATGAAGAA 1080
Db 1021 TGTCCAAAAGCTGAAACAACTCCCAAGTGGTGGTAAAAATTTTGAATTTCTATATGAAGAA 1080
Qy 1081 AATGATGATCTGAGTGGTGGTAAAAATTTTCACTTACATCCCTATGTTGGAAAAACAAGAGG 1140
Db 1081 AATGATGATCTGAGTGGTGGTAAAAATTTTCACTTACATCCCTATGTTGGAAAAACAAGAGG 1140
Qy 1141 GCAGATTCAATCCGTTATCTTAAACAGGTACAGCAATTCCTAAAGTCACTTTTAAGATG 1200
Db 1141 GCAGATTCAATCCGTTATCTTAAACAGGTACAGCAATTCCTAAAGTCACTTTTAAGATG 1200
Qy 1201 TCAGAAATATGATTTGATGCTGAAGAAAGATCATAGATGTCATAGATAAGGAACATAAT 1260
Db 1201 TCAGAAATATGATTTGATGCTGAAGAAAGATCATAGATGTCATAGATAAGGAACATAAT 1260
Qy 1261 CAACCTTTTGAATTTCTATTTCAAGAGTTGAATATATTGCGCAGAGCTGGATGACCTCT 1320
Db 1261 CAACCTTTTGAATTTCTATTTGAAGAGTTGAATATATTGCGCAGAGCTGGATGACCTCT 1320
Qy 1321 GAGGAAAAATATGCTTTGGTCCATCTCTACTAGATGCTCCAGACTCGCCTACAGATAGT 1380
Db 1321 GAGGAAAAATATGCTTTGGTCCATCTCTACTAGATGCTCCAGACTCGCCTACAGATAGT 1380
Qy 1381 TTGATCTCACTGAAATTTTATCCAGTAGAAGATGATTTATGGAAGGCGAGAGCTC 1440
Db 1381 TTGATCTCACTGAAATTTTATCCAGTAGAAGATGATTTATGGAAGGCGAGAGCTC 1440
Qy 1441 ATTGAGTCAGTGGCTGATTTCTGTGAGCCCACTAATTAATCTATGAAGAAACAACAGACTC 1500
Db 1441 ATTGAGTCAGTGGCTGATTTCTGTGAGCCCACTAATTAATCTATGAAGAAACAACAGACTC 1500
Qy 1501 TGGATAAATATCCATGACATCTTTTCAATGTTTTCCTCCCAAGTCAAGAGAGAAATTGAG 1560
Db 1501 TGGATAAATATCCATGACATCTTTTCAATGTTTTCCTCCCAAGTCAAGAGAGAAATTGAG 1560

Db 2161 CTGTTCAATATATGGTGGTC 2179

Search completed: May 4, 2006, 09:35:54
Job time : 15368 secs

Db	1081	AATGATGAATCTGAGGTGAAATATTATCATGTTACATCCCTATGTTGGAACAAGAGG	1140
Qy	1141	GCAGATTCATTCCGTTATCCTTAAACACAGGTACAGCAATCCTTAAAGTCACCTTTTAGATG	1200
Db	1141	GCAGATTCATTCCGTTATCCTTAAACACAGGTACAGCAATCCTTAAAGTCACCTTTTAGATG	1200
Qy	1201	TCAGAAATATGATTCATGCTGAAGGAAGGATCATAGATGTCATAGATAAGGAACCTAAAT	1260
Db	1201	TCAGAAATATGATTCATGCTGAAGGAAGGATCATAGATGTCATAGATAAGGAACCTAAAT	1260
Qy	1261	CAACCTTTTGAGATTCATTTTGAAGGAGTTGAAATATATTGCCAGAGCTGGATGGACTCCT	1320
Db	1261	CAACCTTTTGAGATTCATTTTGAAGGAGTTGAAATATATTGCCAGAGCTGGATGGACTCCT	1320
Qy	1321	GAGGGAAATATGCTTGGTCCATCTTAAGATCGCTCCAGACTCGCCCTACAGATAGTG	1380
Db	1321	GAGGGAAATATGCTTGGTCCATCTTAAGATCGCTCCAGACTCGCCCTACAGATAGTG	1380
Qy	1381	TTGATCTACCTGAATATTATTTCCAGTAGAAGATGATGTTATGGAAGGCAGAGACTC	1440
Db	1381	TTGATCTACCTGAATATTATTTCCAGTAGAAGATGATGTTATGGAAGGCAGAGACTC	1440
Qy	1441	ATTGAGTCAGTCCGTGATCTCTGACGCCACTAATTAATCTATGAAGAAACAACAGACATC	1500
Db	1441	ATTGAGTCAGTCCGTGATCTCTGACGCCACTAATTAATCTATGAAGAAACAACAGACATC	1500
Qy	1501	TGATAAAATATCCATGACATCTTTTCAATGTTTTTCCCAAGTCACGAAGAGGAAATTTGAG	1560
Db	1501	TGATAAAATATCCATGACATCTTTTCAATGTTTTTCCCAAGTCACGAAGAGGAAATTTGAG	1560
Qy	1561	TTTATTTTGGCTCTCTGAATGCAAAAACAGGTTTCGGTCATTTTATACAAAATTAATCATTT	1620
Db	1561	TTTATTTTGGCTCTCTGAATGCAAAAACAGGTTTCGGTCATTTTATACAAAATTAATCATTT	1620
Qy	1621	TTAAAGAAAGCAAAATATAAAGATCCAGTGGTGGGCTGCCTGCTCCAAGTGATTTCAAG	1680
Db	1621	TTAAAGAAAGCAAAATATAAAGATCCAGTGGTGGGCTGCCTGCTCCAAGTGATTTCAAG	1680
Qy	1681	TGTCCTTATCAAAGAGAGATAGCAATTACCAGTGTGTGAATGGGAAGTTCTTGGCCGGCAT	1740
Db	1681	TGTCCTTATCAAAGAGAGATAGCAATTACCAGTGTGTGAATGGGAAGTTCTTGGCCGGCAT	1740
Qy	1741	GGATCTAATATCAAAGTGTGAAGTCAGAAGGCTGGTATATTTTGAAGGCACCAAAAGAC	1800
Db	1741	GGATCTAATATCAAAGTGTGAAGTCAGAAGGCTGGTATATTTTGAAGGCACCAAAAGAC	1800
Qy	1801	TGCCCTTTAGAGCATCACCTGTACGTAGTCAGTTACGTAAATCCTGGAGAGGTGACAAGG	1860
Db	1801	TGCCCTTTAGAGCATCACCTGTACGTAGTCAGTTACGTAAATCCTGGAGAGGTGACAAGG	1860
Qy	1861	CTGACTGACCGTGGCTACTCACATTTCTGCTGCATCAGTCAGCACTGTGACTTCTTTATA	1920
Db	1861	CTGACTGACCGTGGCTACTCACATTTCTGCTGCATCAGTCAGCACTGTGACTTCTTTATA	1920
Qy	1921	AGTAAGTATAGTAACCAAGAAATCCACATCTGTGTCCCTTTTGAAGTACTCAAGTCCT	1980
Db	1921	AGTAAGTATAGTAACCAAGAAATCCACATCTGTGTCCCTTTTGAAGTACTCAAGTCCT	1980
Qy	1981	GAAGATGACCAACTTGCAGAAACAAAGGAATTTTGGGCCACCATTTTGGATTTCAGCAGGT	2040
Db	1981	GAAGATGACCAACTTGCAGAAACAAAGGAATTTTGGGCCACCATTTTGGATTTCAGCAGGT	2040
Qy	2041	CGTCTTCTGACTATATCTCTCCAGAAATTTTCTCTTTTGAAGTACTTCTGGATTTTACA	2100
Db	2041	CGTCTTCTGACTATATCTCTCCAGAAATTTTCTCTTTTGAAGTACTTCTGGATTTTACA	2100
Qy	2101	TTGATGGGATGGCTTACAAAGCCTCATGATCTACAGCCTGGAAAGAAATATCCTACTGTG	2160
Db	2101	TTGATGGGATGGCTTACAAAGCCTCATGATCTACAGCCTGGAAAGAAATATCCTACTGTG	2160
Qy	2161	CTGTTCAATATATGGTGGTC 2179	

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 4, 2006, 04:08:39 ; Search time 1770 Seconds
(without alignments)
11747.933 Million cell updates/sec

Title: US-10-825-632-2
Perfect score: 3120
Sequence: 1 aagtgctaaagctccgagg.....agaattactaaaaaaaaa 3120

Scoring table: OLIGO NUC

Gapop 60.0 , Gapext 60.0

Searched: 4996997 seqs, 3332346308 residues

Wordsize: 6

Total number of hits satisfying chosen parameters: 9993364

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_21.*

- 1: Geneseq1980s.*
- 2: Geneseq1990s.*
- 3: Geneseq2000s.*
- 4: Geneseq2001as.*
- 5: Geneseq2001bs.*
- 6: Geneseq2002as.*
- 7: Geneseq2002bs.*
- 8: Geneseq2003as.*
- 9: Geneseq2003bs.*
- 10: Geneseq2003cs.*
- 11: Geneseq2003ds.*
- 12: Geneseq2004as.*
- 13: Geneseq2004bs.*
- 14: Geneseq2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3120	100.0	3120	4 AAC85694	Aac85694 Nucleotid
2	3120	100.0	3120	6 AAD38956	Aad38956 Human dip
3	3003	96.2	3106	6 ABK12892	Abk12892 Human pro
4	2957	94.8	3143	6 AAH99934	Aah99934 cDNA enco
5	2889	92.6	4829	6 ABK83327	Abk83327 cDNA enco
6	2598	83.3	2649	14 ADVA43981	Adv43981 Human psy
7	2541	81.4	2643	6 AAH99935	Aah99935 Coding se
8	2496	80.0	2649	6 ABX12255	Abx12255 cDNA enco
9	2495	80.0	2671	6 ABK83322	Abk83322 cDNA enco
10	2315	74.2	2797	12 ADL13374	Adl13374 Human ste
11	2305	73.9	2929	10 ACA92425	Aca92425 DNA enco
12	2271	72.8	2952	10 ACA92421	Aca92421 DNA enco
13	2217	71.1	2696	12 ADI16386	Adi16386 Human pro
14	2128	68.2	4523	6 ABK83325	Abk83325 cDNA enco
15	2128	68.2	4676	6 ABK83331	Abk83331 cDNA enco
16	2116	67.8	2842	6 ABN59774	Abn59774 Novel hum
17	1984	63.6	4685	6 ABK83332	Abk83332 cDNA enco
18	1950	62.5	2083	7 ADRA41222	Adra41222 Human CD-
19	1922	61.6	2349	14 ADVA43982	Adv43982 Human psy

20	1770	56.7	1821	6 ABV76411	Abv76411 Dipeptidyl
21	1450	46.5	2161	4 AAH15009	Aah15009 Human cDN
22	1377	44.1	2668	6 ABN59775	Abn59775 Novel hum
23	1191	38.2	4309	6 ABK83328	Abk83328 cDNA enco
24	1164	37.3	2510	6 AAD23843	Aad23843 Human pro
25	1023	32.8	1669	4 AAC85696	Aac85696 Nucleotid
26	956	30.6	2702	13 ADT04072	Adt04072 Human pro
27	877	28.1	1356	6 ABK83326	Abk83326 cDNA enco
28	823	26.4	2251	10 ADE79035	Ade79035 Human pro
29	790	25.3	1197	4 AAC85695	Aac85695 Nucleotid
30	789	25.3	1083	4 AAC85697	Aac85697 Nucleotid
31	760	24.4	832	6 ABK83330	Abk83330 cDNA enco
32	739	23.7	873	4 AAF81719	Aaf81719 Human pro
33	737	23.6	925	6 ABL90148	AbL90148 Human pol
34	654	21.0	2463	10 ACA92424	Aca92424 DNA enco
35	501	16.1	631	4 AAH07860	Aah07860 Human cDN
36	472	15.1	620	6 ABK83329	Abk83329 cDNA enco
37	431	13.8	502	9 ACH32373	Ach32373 Human end
38	309	9.9	823	6 ABK30401	Abk30401 Human G-p
39	254	8.1	561	4 AAL00876	Aal00876 Human rep
40	205	6.6	587	4 AAH12830	Aah12830 Human cDN
41	167	5.4	168	8 ABZ09094	Abz09094 Human oli
42	167	5.4	168	10 ABZ78547	Abz78547 Tumour su
43	156	5.0	4797	4 AAL04386	Aal04386 Human rep
44	82	2.6	89	3 AAC12358	Aac12358 Human sec
45	46	1.5	60	6 ABN42337	Abn42337 Human spl

ALIGNMENTS

RESULT 1
AAC85694
ID AAC85694 standard; cDNA; 3120 BP.
XX
AC AAC85694;
XX
DT 29-JUN-2001 (first entry)
XX
DE Nucleotide sequence of human DPP8.
XX
KW Human; dipeptidyl aminopeptidase; DPP8; prolyl oligopeptidase;
KW dipeptidyl peptidase; DPPIV; T cell; cleavage; diarrhoea;
KW growth hormone deficiency; Glucose level; mucosal regeneration;
KW non-insulin dependent diabetes mellitus; glucose intolerance;
KW immunosuppression; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 214..2862
FT /*tag= a
FT /product= "Human DPP8"
XX
PN WO200119866-A1.
XX
PD 22-MAR-2001.
XX
PF 11-SEP-2000; 2000WO-AU001085.
XX
PR 10-SEP-1999; 99AU-00002762.
PR 18-FEB-2000; 2000AU-00005709.
XX
PA (UNSY) UNIV SYDNEY.
XX
PI Abbott CA, Gorell MD;
DR WPI; 2001-281520/29.
XX
PT P-PSDB; AAB47187.
PT
PT New human dipeptidyl aminopeptidase (DPP8) useful for cleaving substrates, identifying inhibitors of DPP8 catalytic activity which have therapeutic uses, and for detecting activated T cells.

Db 1801 TCCCTTTAGGATCACCTGTAGTACGTAGTAAATCTCGAGAGGTGACAAG 1860
Qy 1861 CTGACTGACCGTGGCTACTCACAATCTTGTGTCATCAGTCAAGCACTGATCTTTATTA 1920
Db 1861 CTGACTGACCGTGGCTACTCACAATCTTGTGTCATCAGTCAAGCACTGATCTTTATTA 1920
Qy 1921 AGTAAGTATAGTAAACAGAGAAATCCACACTGTGTGTCCTTTACAGCTATCAAGTCCT 1980
Db 1921 AGTAAGTATAGTAAACAGAGAAATCCACACTGTGTGTCCTTTACAGCTATCAAGTCCT 1980
Qy 1981 CAAGATGACCAACTCTGCAAAACAAAGGAATTTTGGGCCACCATTTTGGATTACAGCAGGT 2040
Db 1981 GAAGATGACCAACTCTGCAAAACAAAGGAATTTTGGGCCACCATTTTGGATTACAGCAGGT 2040
Qy 2041 CCTCTTCTGATATACCTCTCCAGAAATTTTCTTTTGAAGTACTACTGGATTATCA 2100
Db 2041 CCTCTTCTGATATACCTCTCCAGAAATTTTCTTTTGAAGTACTACTGGATTATCA 2100
Qy 2101 TTGTATGGGATGCTTCAACAGCCTCATGATCTACAGCCTGGAAAGAAATATCTTACTGTG 2160
Db 2101 TTGTATGGGATGCTTCAACAGCCTCATGATCTACAGCCTGGAAAGAAATATCTTACTGTG 2160
Qy 2161 CTGTTTCATATATGTTGCTCTCAGGTGCAAGTGGTGAATATCGGTTTAAAGGAGTCAAG 2220
Db 2161 CTGTTTCATATATGTTGCTCTCAGGTGCAAGTGGTGAATATCGGTTTAAAGGAGTCAAG 2220
Qy 2221 TATTTCCGCTTGATACCTTAGCTCTCTAGTGTATGTTGTTAGTATAGACAAACAGG 2280
Db 2221 TATTTCCGCTTGATACCTTAGCTCTCTAGTGTATGTTGTTAGTATAGACAAACAGG 2280
Qy 2281 GATCTCTGACCCAGAGGCTTAAATTTGAAGGCGCTTTAAATATATAAATGGGTCAATA 2340
Db 2281 GATCTCTGACCCAGAGGCTTAAATTTGAAGGCGCTTTAAATATATAAATGGGTCAATA 2340
Qy 2341 GAAATTTGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTCAATGAC 2400
Db 2341 GAAATTTGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTCAATGAC 2400
Qy 2401 TTAGATCGTGTGGGCATCCACGGCTGGTCTATGAGGATACCTCTCCCTGATGGCATT 2460
Db 2401 TTAGATCGTGTGGGCATCCACGGCTGGTCTATGAGGATACCTCTCCCTGATGGCATT 2460
Qy 2461 ATGAGAGGTGAGATATCTTCAAGGTTGCTTATGCTGGGGCCCCAGTCACTCTGGGATC 2520
Db 2461 ATGAGAGGTGAGATATCTTCAAGGTTGCTTATGCTGGGGCCCCAGTCACTCTGGGATC 2520
Qy 2521 TTCTATGATACAGGATACAGGAACGTTATATGGGTCACTGACAGAAATGAACAGGCG 2580
Db 2521 TTCTATGATACAGGATACAGGAACGTTATATGGGTCACTGACAGAAATGAACAGGCG 2580
Qy 2581 TATTAATTTAGGATCTGTGGCATGCAAGCAGAGAAAGTTCCTCTGAAACCAATCGTTTA 2640
Db 2581 TATTAATTTAGGATCTGTGGCATGCAAGCAGAGAAAGTTCCTCTGAAACCAATCGTTTA 2640
Qy 2641 CTGCTCTTACATGTTTCTGGATGAGAATGTCATTTTGCACATACAGATATATCTG 2700
Db 2641 CTGCTCTTACATGTTTCTGGATGAGAATGTCATTTTGCACATACAGATATATCTG 2700
Qy 2701 AGTTTATTTAGTGGCTGGAAGCCATATGATTTACAGATCTATCTCAGGAGAGAC 2760
Db 2701 AGTTTATTTAGTGGCTGGAAGCCATATGATTTACAGATCTATCTCAGGAGAGAC 2760
Qy 2761 ACATTAAGGTTTCTGATCGGAGAACATTAAGACTGATCTTTTGGCACTACCTTCAA 2820
Db 2761 ACATTAAGGTTTCTGATCGGAGAACATTAAGACTGATCTTTTGGCACTACCTTCAA 2820
Qy 2821 GAAAACTTTGATACAGTATGCTGCTTAAAGTGATATAATTTTGAACCTGTGTAGAAC 2880
Db 2821 GAAAACTTTGATACAGTATGCTGCTTAAAGTGATATAATTTTGAACCTGTGTAGAAC 2880
Qy 2881 TCTCTGGTATACACTGGCTATTTAAACCAATGAGGAGTTTAAATCAACAGAAAAACAGA 2940
Db 2881 TCTCTGGTATACACTGGCTATTTAAACCAATGAGGAGTTTAAATCAACAGAAAAACAGA 2940

Qy 2941 ATTGATCATCACATTTTGTATCTGCGATGTAACATCTACTCTCGAAATAAATGTGGT 3000
Db 2941 ATTGATCATCACATTTTGTATCTGCGATGTAACATCTACTCTCGAAATAAATGTGGT 3000
Qy 3001 CCATGCAAGGGTCTACGGTTTGTGTAGTAAATCTAATACCTTAACCCACATGCTCAAAA 3060
Db 3001 CCATGCAAGGGTCTACGGTTTGTGTAGTAAATCTAATACCTTAACCCACATGCTCAAAA 3060
Qy 3061 TCAATGATACATATTCTCTGAGAGAGCCAGCAATACCAATGATTAATCTAAAAA 3120
Db 3061 TCAATGATACATATTCTCTGAGAGAGCCAGCAATACCAATGATTAATCTAAAAA 3120

RESULT 2
AAD38956
ID AAD38956 standard; cDNA; 3120 BP.
XX AAD38956;
AC AAD38956;
XX
DT 23-SEP-2002 (first entry)
XX
DE Human dipeptidyl peptidase 8 (DPP8) cDNA.
XX
DE Human; dipeptidyl peptidase; DPP; neoplasia; type II diabetes; cirrhosis;
KW autoimmunity; human immuno deficiency virus; HIV infection; cytostatic;
KW graft rejection; antidiabetic; antiinflammatory; immunosuppressive;
KW antiviral; enzyme; gene; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 214..2862
FT /*tag= a
FT /product= "Human DPP8 protein"
XX
PN WO200234900-A1.
XX
XX 02-MAY-2002.
PD
XX
PF 29-OCT-2001; 2001MO-AU001388.
PP
XX
PR 27-OCT-2000; 2000AU-00001078.
XX
XX (UNSY) UNIV SYDNEY.
PA
XX
PI Abbott CA, Gorrell MD;
PI
DR WPI: 2002-454646/48.
DR P-PSDB; AAE24170.
XX
XX New dipeptidyl peptidase (DPP) peptides, useful for screening inhibitors
PT of DPP catalytic activity, which may be employed to treat e.g. neoplasia,
PT type II diabetes, cirrhosis, autoimmunity, graft rejection and HIV
PT infection.
XX
PS Example; Fig 1; 91pp; English.
XX
CC The present invention relates to dipeptidyl peptidase (DPP) proteins and
CC polynucleotides encoding such proteins. The DPP peptides are useful for
CC screening inhibitors of DPP catalytic activity. The inhibitors are useful
CC for treating neoplasia, type II diabetes, cirrhosis, autoimmunity, graft
CC rejection and HIV (human immuno deficiency virus) infection. The present
CC sequence is human DPP8 cDNA
XX
SQ Sequence 3120 BP; 936 A; 637 C; 706 G; 841 T; 0 U; 0 Other;

Query Match 100.0%; Score 3120; DB 6; Length 3120;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAGTCTAAAGCCTCGAGGCAAGCGCTGCTACTGCGCGTCTTCTTAGTGGCG 60
|||||

Db 1 AAGTGTAAAGCCTCCGAGGCCAAGCGCGTGTCTACTGTGCGCGCGTCTTCTTAGTGCCG 60
Qy 61 CGTTCCGCGCGCTGGGTTGTCAACGGCGCGCCGCCGAGGAAGCACTGCAACACGAGACCG 120
Db 61 CGTTCCGCGCGCTGGGTTGTCAACGGCGCGCCGCCGAGGAAGCACTGCAACACGAGACCG 120
Qy 121 GAGTGGAGCGCGCGCAGCATGAGCGCGCGCAGCGCGCTCCATAGCGACGTCGGACCG 180
Db 121 GAGTGGAGCGCGCGCAGCATGAGCGCGCGCAGCGCGCTCCATAGCGACGTCGGACCG 180
Qy 181 TCCGGCGCGCGCGCGGGAAGGAATGCAACATCGCAGCAGCAATGGAACAGACAG 240
Db 181 TCCGGCGCGCGCGCGGGAAGGAATGCAACATCGCAGCAGCAATGGAACAGACAG 240
Qy 241 CTGGGTGTGAGATATTGAACTGCGGACTGTGAGGAGAAATATTGAATCAACAGATCGG 300
Db 241 CTGGGTGTGAGATATTGAACTGCGGACTGTGAGGAGAAATATTGAATCAACAGATCGG 300
Qy 301 CCTAAATTGGAGCCTTTTATGTGTGACGGGTATTCCTGGAGTCAGCTTTAAAAAGCTGCTT 360
Db 301 CCTAAATTGGAGCCTTTTATGTGTGACGGGTATTCCTGGAGTCAGCTTTAAAAAGCTGCTT 360
Qy 361 GCCGATACCGAAAAATATCATGGCTACATGATGGCTTAAGGCCACCAACATGATTTTCATGTTT 420
Db 361 GCCGATACCGAAAAATATCATGGCTACATGATGGCTTAAGGCCACCAACATGATTTTCATGTTT 420
Qy 421 GTGAAGAGGAATGATCCAGATGGACCTCATTCAGACAGAAATCTATTACCTTTGCCATGTCT 480
Db 421 GTGAAGAGGAATGATCCAGATGGACCTCATTCAGACAGAAATCTATTACCTTTGCCATGTCT 480
Qy 481 GGTGAGAACAGAGAAAAATACACTGTGTTTATTTCTGAAATTTCCCAAACTATCAATAGAGCA 540
Db 481 GGTGAGAACAGAGAAAAATACACTGTGTTTATTTCTGAAATTTCCCAAACTATCAATAGAGCA 540
Qy 541 GCAGTCTTAATGTCTCTCTGGAAGCCTCTTTTGGATCTTTTTCAGGCAACACTGGACTAT 600
Db 541 GCAGTCTTAATGTCTCTCTGGAAGCCTCTTTTGGATCTTTTTCAGGCAACACTGGACTAT 600
Qy 601 GGAATGTATTCTCGAGAGAGAACTATTAAAGAGAAAGAAACCAATGAGACAGTCGGA 660
Db 601 GGAATGTATTCTCGAGAGAGAACTATTAAAGAGAAAGAAACCAATGAGACAGTCGGA 660
Qy 661 ATTGCTTTTACGATTTATCACGAAGAGTGAACATTTCTGTTTCAAGCGGTAGTGGGA 720
Db 661 ATTGCTTTTACGATTTATCACGAAGAGTGAACATTTCTGTTTCAAGCGGTAGTGGGA 720
Qy 721 ATTTATCAGGTAAAGATGAGGGCCAAGGATTTACGAAACAACTTTAAAGCCCAAT 780
Db 721 ATTTATCAGGTAAAGATGAGGGCCAAGGATTTACGAAACAACTTTAAAGCCCAAT 780
Qy 781 CTAGTGAAACTAGTTGTGCCCAACATACGAGATGGATCCAAAATTTATGCCCGCTGATCCA 840
Db 781 CTAGTGAAACTAGTTGTGCCCAACATACGAGATGGATCCAAAATTTATGCCCGCTGATCCA 840
Qy 841 GACTGGATTGCTTTTATACATAGCAACGATATTGATATCTAAACATCGTAACACAGAGAA 900
Db 841 GACTGGATTGCTTTTATACATAGCAACGATATTGATATCTAAACATCGTAACACAGAGAA 900
Qy 901 GAAAGGAGACTCACTTATGTGACAAATGAGCTAGCCAACTGGAAGAAAGATGCCAGATCA 960
Db 901 GAAAGGAGACTCACTTATGTGACAAATGAGCTAGCCAACTGGAAGAAAGATGCCAGATCA 960
Qy 961 GCTGGAGTCCCTACCTTTGTTCTCCAAGAAATTTGATAGATATTCGCTATTTGGTGG 1020
Db 961 GCTGGAGTCCCTACCTTTGTTCTCCAAGAAATTTGATAGATATTCGCTATTTGGTGG 1020
Qy 1021 TGTCCAAAAGCTGAACAACTCCAGTGTGTGTAATTTCTTAGAATTTCTATATGAAGAA 1080
Db 1021 TGTCCAAAAGCTGAACAACTCCAGTGTGTGTAATTTCTTAGAATTTCTATATGAAGAA 1080
Qy 1081 AATGATGAATCTGAGGTGAATATTATCTATCATCCCTATGTTGGAAACAGAGG 1140
Db 1081 AATGATGAATCTGAGGTGAATATTATCTATCATCCCTATGTTGGAAACAGAGG 1140

Qy 1141 GCAGATTCATTCGGTTATCTTAATAACAGGTACAGCAATCTCTAAAGTCACCTTTTAAGATG 1200
Db 1141 GCAGATTCATTCGGTTATCTTAATAACAGGTACAGCAATCTCTAAAGTCACCTTTTAAGATG 1200
Qy 1201 TCAGAAATATGATTTGATGCTGAAGGAGGATCATAGATGTCATAGATAAGAACTAAAT 1260
Db 1201 TCAGAAATATGATTTGATGCTGAAGGAGGATCATAGATGTCATAGATAAGAACTAAAT 1260
Qy 1261 CAACCTTTTGAGATTTCTATTTGAAGGAGTTGAATATTATGCGAGAGCTGGATGGAATCCT 1320
Db 1261 CAACCTTTTGAGATTTCTATTTGAAGGAGTTGAATATTATGCGAGAGCTGGATGGAATCCT 1320
Qy 1321 GAGGAAAAATATGCTTTGGTCCATCTCTAGATCGCTCCAGACTCGCTTACAGATAGTG 1380
Db 1321 GAGGAAAAATATGCTTTGGTCCATCTCTAGATCGCTCCAGACTCGCTTACAGATAGTG 1380
Qy 1381 TTGATCTCACCTGAAATTTATTTATCCAGTAGAAGATGATGTTATGGAAGGCAGAGACTC 1440
Db 1381 TTGATCTCACCTGAAATTTATTTATCCAGTAGAAGATGATGTTATGGAAGGCAGAGACTC 1440
Qy 1441 ATTGAGTCAGTGCCTGATTTCTGTGACGCCACTAAATTTATCTATGAAGAAACACAGACATC 1500
Db 1441 ATTGAGTCAGTGCCTGATTTCTGTGACGCCACTAAATTTATCTATGAAGAAACACAGACATC 1500
Qy 1501 TGGATAAATATCCATGACATCTTTTCATGTTTTTCCCAAGTCACGAGAGGAAATTTGAG 1560
Db 1501 TGGATAAATATCCATGACATCTTTTCATGTTTTTCCCAAGTCACGAGAGGAAATTTGAG 1560
Qy 1561 TTTATTTTGGCTCTCAATGCAAAACAGGTTTCGGTCATTTTACAAATTTACATCTATT 1620
Db 1561 TTTATTTTGGCTCTCAATGCAAAACAGGTTTCGGTCATTTTACAAATTTACATCTATT 1620
Qy 1621 TTAAGGAAAGCAAAATATAACGATCCAGTGGTGGGCTGCTGCTCCAGTGATTTCAAG 1680
Db 1621 TTAAGGAAAGCAAAATATAACGATCCAGTGGTGGGCTGCTGCTCCAGTGATTTCAAG 1680
Qy 1681 TGCTCTATCAAGAGAGATAGCAATTTACAGTGGTGAATGGGAAGTTCTTTGGCCGGCAT 1740
Db 1681 TGCTCTATCAAGAGAGATAGCAATTTACAGTGGTGAATGGGAAGTTCTTTGGCCGGCAT 1740
Qy 1741 GGATCTAATACTCAAGTTGATGAAGTCAAGAGGCTGGTATATTTTGAAGGCACCAAGAC 1800
Db 1741 GGATCTAATACTCAAGTTGATGAAGTCAAGAGGCTGGTATATTTTGAAGGCACCAAGAC 1800
Qy 1801 TCCCTTTTAGAGCATCACCTGTACGTAGTTCAGTTACGTAAATCTCTGGAGAGGTGACAAGG 1860
Db 1801 TCCCTTTTAGAGCATCACCTGTACGTAGTTCAGTTACGTAAATCTCTGGAGAGGTGACAAGG 1860
Qy 1861 CTGACTGACCGTGGCTACTCACAATTTCTGCTGCATCAGTCAGCACTGTGACTTCTTTATA 1920
Db 1861 CTGACTGACCGTGGCTACTCACAATTTCTGCTGCATCAGTCAGCACTGTGACTTCTTTATA 1920
Qy 1921 AGTAAGTATAGTAAACAGAAAGATCCACATGTGTGCTCCCTTTTACAAGCTATCAAGTCCT 1980
Db 1921 AGTAAGTATAGTAAACAGAAAGATCCACATGTGTGCTCCCTTTTACAAGCTATCAAGTCCT 1980
Qy 1981 GAAGATGACCCCAACTTGGCAAAACAAAGAAATTTTGGGCCCACTTTTGGATTCAGAGGT 2040
Db 1981 GAAGATGACCCCAACTTGGCAAAACAAAGAAATTTTGGGCCCACTTTTGGATTCAGAGGT 2040
Qy 2041 CCTCTTCTGACTATCTCTCCAGAAATTTCTCTTTTGAAGTACTACTGATTTTACA 2100
Db 2041 CCTCTTCTGACTATCTCTCCAGAAATTTCTCTTTTGAAGTACTACTGATTTTACA 2100
Qy 2101 TTGTATGGGATGCTCTACAGCTCATGATCTCAGCTCGAAGAAATATCTCTACTGTG 2160
Db 2101 TTGTATGGGATGCTCTCAGCTCATGATCTCAGCTCGAAGAAATATCTCTACTGTG 2160
Qy 2161 CTGTTCAATATATGTTGGTCTCTCAGGTGAGTTGTTGAATATCTCGTTTAAAGGAGTCAAG 2220
Db 2161 CTGTTCAATATATGTTGGTCTCTCAGGTGAGTTGTTGAATATCTCGTTTAAAGGAGTCAAG 2220

QY 2221 TATTTCGGCTTGAATACCTTAGCTCTCTAGCTTATGTGGTTAGTATAGATAGACACAGG 2280
DB |||||
QY 2221 TATTTCGGCTTGAATACCTTAGCTCTCTAGCTTATGTGGTTAGTATAGATAGACACAGG 2280
DB |||||
QY 2281 GATCTCTGTACCGAGGGCTTAAATTTGAAGGGCTTTAAATATATAAATGGGTCAATA 2340
DB GATCTCTGTACCGAGGGCTTAAATTTGAAGGGCTTTAAATATATAAATGGGTCAATA 2340
QY 2341 GAAATTCACGATCAGGTGGAAGGACTCCAAATATCTAGCTTCTCGATATGATTTTCATTGAC 2400
DB GAAATTCACGATCAGGTGGAAGGACTCCAAATATCTAGCTTCTCGATATGATTTTCATTGAC 2400
QY 2401 TTATGATCGTGTGGGCATCCACGGCTGTCTCTATGAGGATACCTCTCCCTGATGGCATT 2460
DB TTATGATCGTGTGGGCATCCACGGCTGTCTCTATGAGGATACCTCTCCCTGATGGCATT 2460
QY 2461 ATGACAGGTTCAGATATCTTCAGGGTGTCTATTTGCTGGGGCCCGACGTCTCTGTGATC 2520
DB ATGACAGGTTCAGATATCTTCAGGGTGTCTATTTGCTGGGGCCCGACGTCTCTGTGATC 2520
QY 2521 TTCTATGATACAGGATACAGGAACTTATATGAGGTCACCTTGACCAAGATGAACAGGC 2580
DB TTCTATGATACAGGATACAGGAACTTATATGAGGTCACCTTGACCAAGATGAACAGGC 2580
QY 2581 TATTACTAGGATCTGTGGCATGCAAGCAGAAAAGTTCCCTCTGAAACCAAAATCGTTTA 2640
DB TATTACTAGGATCTGTGGCATGCAAGCAGAAAAGTTCCCTCTGAAACCAAAATCGTTTA 2640
QY 2641 CTGCTCTTACATGTTTCTGAGATGAGAAATGTCATTTTGCACATACACAGTATATTAATG 2700
DB CTGCTCTTACATGTTTCTGAGATGAGAAATGTCATTTTGCACATACACAGTATATTAATG 2700
QY 2701 AGTTTTTTTGTAGGGCTGGAAGCCATATGATTTTACAGATCTATCTCCAGAGAGACAC 2760
DB AGTTTTTTTGTAGGGCTGGAAGCCATATGATTTTACAGATCTATCTCCAGAGAGACAC 2760
QY 2761 AGCATAAGAGTTCTCGAATCGGAGAACATTATGAACTGCACTTTTGCATCTTCAA 2820
DB AGCATAAGAGTTCTCGAATCGGAGAACATTATGAACTGCACTTTTGCATCTTCAA 2820
QY 2821 GAAACCTTGATACGATATGCTGCTCTTAAAGTGATATATTTTGGACCTGTGTAGAAC 2880
DB GAAACCTTGATACGATATGCTGCTCTTAAAGTGATATATTTTGGACCTGTGTAGAAC 2880
QY 2881 TCTCTGTATACACTGCTATTTAAACCAATGAGGAGTTTAAATCAACAGAAAACACAGA 2940
DB TCTCTGTATACACTGCTATTTAAACCAATGAGGAGTTTAAATCAACAGAAAACACAGA 2940
QY 2941 ATTGATCATCATTTTGTATACCTGCGATGATCACTACTCTCTGAAAATAAATGTGGTG 3000
DB ATTGATCATCATTTTGTATACCTGCGATGATCACTACTCTCTGAAAATAAATGTGGTG 3000
QY 3001 CATTGACAGGGTCTACGGTTTGTGGTAGTAATCTAATACCTTAAACCCACATGCTCAAAA 3060
DB CATTGACAGGGTCTACGGTTTGTGGTAGTAATCTAATACCTTAAACCCACATGCTCAAAA 3060
QY 3061 TCAATGATATATTTCTGAGAGACCCAGCAATACCAATGATTAATAAAAAA 3120
DB TCAATGATATATTTCTGAGAGACCCAGCAATACCAATGATTAATAAAAAA 3120

RESULT 3
ID ABK12892
XX
AC ABK12892;
DT 09-APR-2002 (first entry)
XX
DE Human protease PRPS-9 cDNA sequence.
KW Human; protease; PRPS; gastrointestinal; Crohn's disease; cancer;
KW cardiovascular; atherosclerosis; autoimmune disorder; dermatitis;

KW inflammatory disorder; acquired immunodeficiency syndrome; AIDS;
KW cell proliferative disorder; developmental disorder; epilepsy;
KW Duchenne muscular dystrophy; epithelial disorder; neurological disorder;
KW reproductive disorder; endometriosis; ss.
OS Homo sapiens.
FH Key Location/Qualifiers
XX CDS 203..2851
FT /*tag= a
FT /product= "Human protease PRPS-9"
XX
PN WO200198468-A2.
XX
PD 27-DEC-2001.
XX
PF 13-JUN-2001; 2001WO-US019178.
PR 16-JUN-2000; 2000US-0212336P.
PR 22-JUN-2000; 2000US-0213955P.
PR 29-JUN-2000; 2000US-0215396P.
PR 07-JUL-2000; 2000US-0216821P.
PR 14-JUL-2000; 2000US-0218946P.
XX (INCY-) INCYTE GENOMICS INC.
XX Yue H, Elliott VS, Gandhi AR, Lal P, Au-Young J, Tribouley CM;
PI Deleane AM, Baughn MR, Nguyen DB, Lee EA, Hafalia A, Khan FA;
PI Waila NK, Yao MG, Lu DAM, Patterson C, Tang YT, Walsh RT;
PI Azimzai Y, Lu Y, Ramkumar J, Xu Y, Reddy R, Das D, Kearney L;
PI Kallick DA;
XX WPI; 2002-090437/12.
DR P-PSDB; AAU74749.
XX
PT Twenty one human proteases (referred to as PRPS-1 to PRPS-21), useful in
PT the diagnosis, treatment and prevention of gastrointestinal (e.g.
PT gastritis), cardiovascular (e.g. atherosclerosis) and cell proliferative
PT (e.g. cancer) disorders.
XX
PS Claim 5; Page 166-167; 177pp; English;
XX
CC The present invention relates to twenty one new human proteases, referred
CC to as PRPS-1 to PRPS-21. The PRPS polynucleotides and polypeptides of the
CC invention are useful in the diagnosis, treatment and prevention of
CC gastrointestinal e.g. gastritis, esophageal carcinoma and Crohn's
CC disease, cardiovascular e.g. atherosclerosis, hypertension and myocardial
CC infarction, autoimmune/inflammatory e.g. acquired immunodeficiency
CC syndrome (AIDS), allergies and osteoarthritis, cell proliferative e.g.
CC cancer, developmental e.g. Duchenne and Becker muscular dystrophy,
CC epithelial e.g. dermatitis, neurological e.g. epilepsy and Alzheimer's
CC disease and reproductive e.g. infertility and endometriosis disorders.
CC Numerous other examples of each disorder are given in the specification.
CC The present nucleic acid sequence encodes the human protease PRPS-9
CC protein of the invention
XX
SQ Sequence 3106 BP; 928 A; 633 C; 704 G; 841 T; 0 U; 0 Other;

Query Match 96.2%; Score 3003; DB 6; Length 3106;
Best Local Similarity 99.9%; Pred. No; 0;
Matches 3103; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 13 CTCGAGGCCAAGGCCGCTCTACTGCGCGCGCTGCTTCTTAGTCCCGCTTCGCGCCT 72
DB |||||
QY 2 CTCGAGGCCAAGGCCGCTCTACTGCGCGCGCTGCTTCTTAGTCCCGCTTCGCGCCT 61
DB |||||
QY 73 GGGTTGTACCGCGCGCGCGCGCGCGCTGCTTCTTAGTCCCGCTTCGCGCCT 132
DB |||||
QY 62 GGGTTGTACCGCGCGCGCGCGCGCGCTGCTTCTTAGTCCCGCTTCGCGCCT 121
DB |||||
QY 133 CGCAGCATGAGCGCGCGCGCGCGCGCTTCCATAGCGCACGCTCGGAGCGGCGG 192
DB |||||
QY 122 CGCAGCATGAGCGCGCGCGCGCGCGCTTCCATAGCGCACGCTCGGAGCGGCGG 181
DB |||||

QY 193 CGGGGGAAGAAAATCAACATGGCAGCAGCAATGGAAACAGAACACAGCTCGGTGTTGAG 252
Db 182 CGGGGGAAGAAAATCAACATGGCAGCAGCAATGGAAACAGAACACAGCTCGGTGTTGAG 241
QY 253 ATATTTGAAACTGCGGACTGTGAGGAGATATTGAATCACAGAGTCGGCTAAATTTGAG 312
Db 242 ATATTTGAAACTGCGGACTGTGAGGAGATATTGAATCACAGAGTCGGCTAAATTTGAG 301
QY 313 CCTTTTATGTTGAGCGGTATCTCGAGTCAGCTTTAAAGACTGCTTTGCCGATACCAGA 372
Db 302 CCTTTTATGTTGAGCGGTATCTCGAGTCAGCTTTAAAGACTGCTTTGCCGATACCAGA 361
QY 373 AAATATCATGGCTACATGATGGCTAAGGACACACATGATTTCAATGTTTGTGAAGGAAAT 432
Db 362 AAATATCATGGCTACATGATGGCTAAGGACACACATGATTTCAATGTTTGTGAAGGAAAT 421
QY 433 GATCCAGATGACCTCATTCAGACAGAAATCTATTACCTTGGCATGTCTGGTGAGAACAGA 492
Db 422 GATCCAGATGACCTCATTCAGACAGAAATCTATTACCTTGGCATGTCTGGTGAGAACAGA 481
QY 493 GAAATACACTGTTTATCTGAAATTTCCAAAACCTATCAATAGACGACAGTCTTAAATG 552
Db 482 GAAATACACTGTTTATCTGAAATTTCCAAAACCTATCAATAGACGACAGTCTTAAATG 541
QY 553 CTCTCTTGGAGCCTCTTTTGGATCTTTTTCAGGCAACACTGGACTATGGAATGTATCT 612
Db 542 CTCTCTTGGAGCCTCTTTTGGATCTTTTTCAGGCAACACTGGACTATGGAATGTATCT 601
QY 613 CGAGAAGAAACTATTAGAGAAAGAAAACCATTTGGACAGTCGGAATTTGCTTCTTAC 672
Db 602 CGAGAAGAAACTATTAGAGAAAGAAAACCATTTGGACAGTCGGAATTTGCTTCTTAC 661
QY 673 GATTATCACCAAGGAGTGGACATTTCTGTTTTCAGCGGTAGTGGAAATTTATACGTA 732
Db 662 GATTATCACCAAGGAGTGGACATTTCTGTTTTCAGCGGTAGTGGAAATTTATACGTA 721
QY 733 AAAGATGGAGGCCACAGAGNTTACGCAACCTTTAAGCCCAATCTAGTGGAACCT 792
Db 722 AAAGATGGAGGCCACAGAGNTTACGCAACCTTTAAGCCCAATCTAGTGGAACCT 781
QY 793 AGTTGTCGCAACATACGATGATCCAAAATTTATGCCCGCTGATCCAGACTGGATGCT 852
Db 782 AGTTGTCGCAACATACGATGATCCAAAATTTATGCCCGCTGATCCAGACTGGATGCT 841
QY 853 TTTATACATAGCAACGATATTTGGATATCTAACATCGTAAACAGAGAAAGGAGACTC 912
Db 842 TTTATACATAGCAACGATATTTGGATATCTAACATCGTAAACAGAGAAAGGAGACTC 901
QY 913 ACTATGTGCACATGAGCTAGCCAAACATGGAAGAAGATGCCAGATCAGCTGGAGTCGCT 972
Db 902 ACTATGTGCACATGAGCTAGCCAAACATGGAAGAAGATGCCAGATCAGCTGGAGTCGCT 961
QY 973 ACCTTTGTTCTCCAAGAAAGATTTGATAGATATTCTGGCTATTGGTGTCCAAAAGCT 1032
Db 962 ACCTTTGTTCTCCAAGAAAGATTTGATAGATATTCTGGCTATTGGTGTCCAAAAGCT 1021
QY 1033 GAAACAACTCCCAAGTGGTGAATTTCTTAGAATTTCTATAGAAATATGATGATCT 1092
Db 1022 GAAACAACTCCCAAGTGGTGAATTTCTTAGAATTTCTATAGAAATATGATGATCT 1081
QY 1093 GAGGTGGAATTTATCATGTTACATCCCTATGTTGGAAACNAGGAGGCGAGATTCATTC 1152
Db 1082 GAGGTGGAATTTATCATGTTACATCCCTATGTTGGAAACNAGGAGGCGAGATTCATTC 1141
QY 1153 CGTTATCTTAAACAGGTACAGCAATCTTAAAGTCACTTTTAAAGATGTCAGAAATATG 1212
Db 1142 CGTTATCTTAAACAGGTACAGCAATCTTAAAGTCACTTTTAAAGATGTCAGAAATATG 1201
QY 1213 ATGTAGTGTGAAGGAGATCATAGATGTATAGATAAGGAATTAATCAACTTTTGGAG 1272
Db 1202 ATGTAGTGTGAAGGAGATCATAGATGTATAGATAAGGAATTAATCAACTTTTGGAG 1261

QY 1273 ATTCTATTTGAAGAGTTGAATATATATGCGCAGAGCTGGATGAGCTCCTGAAGGAAAAATAT 1332
Db 1262 ATTCTATTTGAAGAGTTGAATATATATGCGCAGAGCTGGATGAGCTCCTGAAGGAAAAATAT 1321
QY 1333 GCTTGGTCCATCTCTACTAGATCGCTCCCAGACTCGCCTACAGATAGTGTGATCTCACCT 1392
Db 1322 GCTTGGTCCATCTCTACTAGATCGCTCCCAGACTCGCCTACAGATAGTGTGATCTCACCT 1381
QY 1393 GAAATATTTATCCCAAGTGAAGATGATGTTATGGAAGGCGAGAGACTCATTTGAGTCAGTG 1452
Db 1382 GAAATATTTATCCCAAGTGAAGATGATGTTATGGAAGGCGAGAGACTCATTTGAGTCAGTG 1441
QY 1453 CCTGATTTCTGTGACGCCACTAATTTATCTATGAAGAAAACAAACAGACATCTGGATAAATATC 1512
Db 1442 CCTGATTTCTGTGACGCCACTAATTTATCTATGAAGAAAACAAACAGACATCTGGATAAATATC 1501
QY 1513 CATGACATCTTTTCAATGTTTTTCCCCAAAAGTCACGAAGAGGAAATTTGAGTTTTTTTTGCC 1572
Db 1502 CATGACATCTTTTCAATGTTTTTCCCCAAAAGTCACGAAGAGGAAATTTGAGTTTTTTTTGCC 1561
QY 1573 TCTGAATGCAAAACAGGTTTTCCGTCAATTTATACAAAATTTACATCTATTTTAAAGGAAGC 1632
Db 1562 TCTGAATGCAAAACAGGTTTTCCGTCAATTTATACAAAATTTACATCTATTTTAAAGGAAGC 1621
QY 1633 AAATATAAACGATCCAGTCGGTGGCTGCTGCTCCAGTGATTTCAAGTGTCCTATCAAA 1692
Db 1622 AAATATAAACGATCCAGTCGGTGGCTGCTGCTCCAGTGATTTCAAGTGTCCTATCAAA 1681
QY 1693 GAGGAGTAGCAATTTACCAAGTGGTGGTGAATGGGAAGTTCTTGGCCGGCATGGAATTAATC 1752
Db 1682 GAGGAGTAGCAATTTACCAAGTGGTGGTGAATGGGAAGTTCTTGGCCGGCATGGAATTAATC 1741
QY 1753 CAAGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGCACCAAGACTCCCTTTTAGAG 1812
Db 1742 CAAGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGCACCAAGACTCCCTTTTAGAG 1801
QY 1813 CATCACTGTGACTAGTCAGTTACGTAATCTCTGGAGAGGTGACAAGGCTGACTGACCGT 1872
Db 1802 CATCACTGTGACTAGTCAGTTACGTAATCTCTGGAGAGGTGACAAGGCTGACTGACCGT 1861
QY 1873 GGCTACTCACTTTCTGTGTCATCAGTCAGCACTGTGACTTCTTTTATAAGTAAAGTATAGT 1932
Db 1862 GGCTACTCACTTTCTGTGTCATCAGTCAGCACTGTGACTTCTTTTATAAGTAAAGTATAGT 1921
QY 1933 AACCAAGAATCCACACTGTGTGTCCTTTTACAAGCTATCAAGTCTCTGAAGATGACCCA 1992
Db 1922 AACCAAGAATCCACACTGTGTGTCCTTTTACAAGCTATCAAGTCTCTGAAGATGACCCA 1981
QY 1993 ACTTGCAAAACAAAGGAATTTTGGGCCACCACTTTTGGATTACAGAGTCTCTTCTTGAC 2052
Db 1982 ACTTGCAAAACAAAGGAATTTTGGGCCACCACTTTTGGATTACAGAGTCTCTTCTTGAC 2041
QY 2053 TATACTCTCCAGAAATTTCTCTTTTGAAGTACTACTGGATTTACATTTGATGGATG 2112
Db 2042 TATACTCTCCAGAAATTTCTCTTTTGAAGTACTACTGGATTTACATTTGATGGATG 2101
QY 2113 CTCTACAGCCCTCATGATCTACAGCTGGAAAGAAATATCTTACTGTGCTGTTTATATAT 2172
Db 2102 CTCTACAGCCCTCATGATCTACAGCTGGAAAGAAATATCTTACTGTGCTGTTTATATAT 2161
QY 2173 GGTGTCCTCAGGTCAGTGGTGAATAAATTCGGTTTAAAGGAGTCAAGTATTTCCGCTTG 2232
Db 2162 GGTGTCCTCAGGTCAGTGGTGAATAAATTCGGTTTAAAGGAGTCAAGTATTTCCGCTTG 2221
QY 2233 AATACCTTAGCCTCTCTAGGTTATGTGGTTGTAGTGATAGACAACAGGGGATCCTGTAC 2292
Db 2222 AATACCTTAGCCTCTCTAGGTTATGTGGTTGTAGTGATAGACAACAGGGGATCCTGTAC 2281
QY 2293 CGAGGCTTAAATTTGAAAGGCGCTTTTAAATATAAAATGGGTCAAAATAGAAAATTTGACGAT 2352
Db 2282 CGAGGCTTAAATTTGAAAGGCGCTTTTAAATATAAAATGGGTCAAAATAGAAAATTTGACGAT 2341
QY 2353 CAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTTCACTTAGATCGTGTG 2412

Db 2342 CAGGTGAGAGCACTCAATATCTAGCTTCGATATGATTTTCATTGACTTAGATCGGTG 2401
QY 2413 GGATCCACGGCTGGTCTTATGGAGATACCTCCCTGATGGCATTAATGCGAGGTCA 2472
Db 2402 GGCATCCACGGCTGGTCTTATGGAGATACCTCCCTGATGGCATTAATGCGAGGTCA 2461
QY 2473 GATATCTTCAAGGCTGGTCTTATGGAGATACCTCCCTGATGGCATTAATGCGAGGTCA 2532
Db 2462 GATATCTTCAAGGCTGGTCTTATGGAGATACCTCCCTGATGGCATTAATGCGAGGTCA 2521
QY 2533 GGATACACGGAACTGATATGGGTCACTCCCTGATGGCATTAATGCGAGGTCA 2592
Db 2522 GGATACACGGAACTGATATGGGTCACTCCCTGATGGCATTAATGCGAGGTCA 2581
QY 2593 TCTGTGGCATGCAAGCAAGATGTTCCCTGATGGCATTAATGCGAGGTCA 2652
Db 2582 TCTGTGGCATGCAAGCAAGATGTTCCCTGATGGCATTAATGCGAGGTCA 2641
QY 2653 GGTTCCTGATGAGATGTTCCATTTTGACATACAGTATATTAAGTCTTCTAGT 2712
Db 2642 GGTTCCTGATGAGATGTTCCATTTTGACATACAGTATATTAAGTCTTCTAGT 2701
QY 2713 AGGCTGGAAAGCCATATGATTTACAGATCTATCTCAGGAGACACAGCATAGAGTT 2772
Db 2702 AGGCTGGAAAGCCATATGATTTACAGATCTATCTCAGGAGACACAGCATAGAGTT 2761
QY 2773 CCTGAATCGGGAGAACATTAATGATGATCTTTTGCATCTCTTCAAGAAACCTTGG 2832
Db 2762 CCTGAATCGGGAGAACATTAATGATGATCTTTTGCATCTCTTCAAGAAACCTTGG 2821
QY 2833 TCAGTATCTGCTCTTAAAGTGATATAATTTTGACCTGTGTAGAACTCTCTGGTATAC 2892
Db 2822 TCAGTATCTGCTCTTAAAGTGATATAATTTTGACCTGTGTAGAACTCTCTGGTATAC 2881
QY 2893 ACTGGCTATTAAACCAATGAGAGGTTTAATCAACAGAAACACAGAAATGATCATC 2952
Db 2882 ACTGGCTATTAAACCAATGAGAGGTTTAATCAACAGAAACACAGAAATGATCATC 2941
QY 2953 ATTTTGATCTGCTGATTAACATCTACTCTGAAATAATGTTGGTCCATGAGGGT 3012
Db 2942 ATTTTGATCTGCTGATTAACATCTACTCTGAAATAATGTTGGTCCATGAGGGT 3001
QY 3013 CTACGGTTTGTGGTAGTAATCTAATACCTTAACCCCATGCTCAAAATCAAAATGATACA 3072
Db 3002 CTACGGTTTGTGGTAGTAATCTAATACCTTAACCCCATGCTCAAAATCAAAATGATACA 3061
QY 3073 TATTCCTGAGACCCAGCAATACCATAGAAATTAATAAAAAA 3117
Db 3062 TATTCCTGAGACCCAGCAATACCATAGAAATTAATAAAAAA 3106

RESULT 4
AAH99934 standard; cDNA; 3143 BP.

XX AC AAH99934;

XX DT 12-APR-2002 (first entry)

XX DE cDNA encoding 21953 human prollyl oligopeptidase.

XX KW 21953 prollyl oligopeptidase; human; proline; endopeptidase; cancer;
XX KW cardiovascular disease; autoimmune disease; atopic allergy;
XX KW neuronal disorder; vascular disease; prostate disorder; cytostatic;
XX KW antidiabetic; antiarthritic; antiasthmatic; antiinflammatory;
XX KW diabetes mellitus; arthritis; multiple sclerosis; asthma;
XX KW Grave's disease; neuronal disorder; demyelinating disease; ss.

XX OS Homo sapiens.

XX Key Location/Qualifiers
FH 229. .2877
FT CDS

FT FT /*tag= a
FT FT /product= "21953 prollyl oligopeptidase"
FT FT /note= "This region is specifically claimed in claim 2"
XX PN WO200179473-A2.
XX PD 25-OCT-2001.
XX PF 11-APR-2001; 2001WO-US040483.
XX PR 18-APR-2000; 2000US-0197508P.
XX PA (MILL-) MILLENNIUM PHARM INC.
XX PI Meyers RA, Williamson M;
XX WPI: 2002-034353/04.
XX P-PSDB; AAG78415.
XX New polypeptides 21953, member of human prollyl oligopeptidase family,
XX useful as diagnostic targets and therapeutic agents for controlling
XX cancer, lymphoma and leukemia.
XX Claim 7; Page 100-102; 121pp; English.
XX This invention relates to an isolated 21953 human prollyl oligopeptidase.
XX Which is cytostatic, antidiabetic, antiarthritic, neuroprotective,
XX antithyroid, dermatological, antipsoriatic, antiasthmatic,
XX ophthalmological, antiinflammatory, nootropic, antianginal, cardiac,
XX anticonvulsant, gynaecological, vasotropic, antianginal, cardiac,
XX antiatherosclerotic, anorectic and metabolic in its action. Uses include
XX gene therapy, expression or activity of 21953 protein modulator, it is
XX useful for identifying a compound which binds to it and can be used in
XX preventing, treating or detecting a cellular proliferative or
XX differentiative disorder. The 21953 molecules can act as novel diagnostic
XX targets and therapeutic agents for controlling disorders associated with
XX associated with cell differentiation and proliferation such as cancer,
XX immune function, reproductive, neurological and cardiovascular function.
XX The 21953 molecules are thus useful for treating and preventing cellular
XX proliferative and differentiative disorders, haematopoietic neoplastic
XX disorders, immune disorders such as autoimmune diseases, diabetes
XX mellitus, arthritis, multiple sclerosis, asthma, Grave's disease,
XX neuronal disorders, demyelinating diseases, vascular disorders and
XX metabolism or pain disorders. This sequence represents the cDNA encoding
XX sequence of 21953 human prollyl oligopeptidase

SQ Sequence 3143 BP; 943 A; 644 C; 712 G; 844 T; 0 U; 0 Other;

Query Match 94.8%; Score 2987; DB 6; Length 3143;
Best Local Similarity 99.9%; Pred. No; 0;
Matches 3107; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 11 GCCTCCGAGGCCAAGCGCGTCTACTGCGCGCGTCTTCTTAGTGGCGGTTCCGCCG 70
Db 26 GCCTCCGAGGCCAAGCGCGTCTACTGCGCGCGTCTTCTTAGTGGCGGTTCCGCCG 85
QY 71 CTGGGTTGTCAACCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 130
Db 86 CTGGGTTGTCAACCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 145
QY 131 GGCGCAGCATGAACCG 190
Db 146 GGCGCAGCATGAACCG 205
QY 191 GCCCGGGGAGGAAATGCAACATGCGAGCAGCAATGAAACAGAACAGCTGGGTGTG 250
Db 206 GCCCGGGGAGGAAATGCAACATGCGAGCAGCAATGAAACAGAACAGCTGGGTGTG 265
QY 251 AGATATTTGAAACTGCGGACTGTGAGGAGAAATTAATGAATCAAGATCGGCCTAAATGG 310
Db 266 AGATATTTGAAACTGCGGACTGTGAGGAGAAATTAATGAATCAAGATCGGCCTAAATGG 325

QY 311 AGCCTTTTATGTTGAGCGGTATTCTTGGAGTCAGCTTAAAGAGCTGCTTGCCTATCCA 370
Db 326 AGCCTTTTATGTTGAGCGGTATTCTTGGAGTCAGCTTAAAGAGCTGCTTGCCTATCCA 385
QY 371 GAAATATATCATGCGCTACATGATGCTAAGGCACCATCATGATTTTCAATGTTTGTGAAGAGGA 430
Db 386 GAAATATATCATGCGCTACATGATGCTAAGGCACCATCATGATTTTCAATGTTTGTGAAGAGGA 445
QY 431 ATGATCCAGATGGAGCTTCAATTCAGACAGAAATCTATTACCTTGCCTATGCTCGTGAGAAC 490
Db 446 ATGATCCAGATGGAGCTTCAATTCAGACAGAAATCTATTACCTTGCCTATGCTCGTGAGAAC 505
QY 491 GAGAAATATACACTGTTTATTTCTGAAATCCCAAACTATCAATAGACAGCACTCTTAA 550
Db 506 GAGAAATATACACTGTTTATTTCTGAAATCCCAAACTATCAATAGACAGCACTCTTAA 565
QY 551 TGCTCTCTTGGAGCGCTTTTGGATCTTTTTCAGGCAACACTGACATATGGAATGATT 610
Db 566 TGCTCTCTTGGAGCGCTTTTGGATCTTTTTCAGGCAACACTGACATATGGAATGATT 625
QY 611 CTCGAGAGAGAACTATTAAAGAGAAAGAAACGCAATTTGGAACAGTCGGAATTCCTTCTT 670
Db 626 CTCGAGAGAGAACTATTAAAGAGAAAGAAACGCAATTTGGAACAGTCGGAATTCCTTCTT 685
QY 671 ACGATTATACCAAGGAAGTGGAACTTTCTGTTTCAAGCCGGTAGTGGAAATTTATCACG 730
Db 686 ACGATTATACCAAGGAAGTGGAACTTTCTGTTTCAAGCCGGTAGTGGAAATTTATCACG 745
QY 731 TAAAGATGGAGGCGCACAGGATTTACGAAACAACTTTAAGGCCCAATCTAGTGGAAA 790
Db 746 TAAAGATGGAGGCGCACAGGATTTACGAAACAACTTTAAGGCCCAATCTAGTGGAAA 805
QY 791 CTAGTTGTCCTCAACATACGAGTGGATCCAAATTTATGCCCCTGATCCAGACTGGATTG 850
Db 806 CTAGTTGTCCTCAACATACGAGTGGATCCAAATTTATGCCCCTGATCCAGACTGGATTG 865
QY 851 CTTTATATACATAGCAACGATATTGATATCTTAAATCTGTAACAGAGAAAGAGAGAC 910
Db 866 CTTTATATACATAGCAACGATATTGATATCTTAAATCTGTAACAGAGAAAGAGAGAC 925
QY 911 TCATTTATGTCACAAATAGCTAGCCAAATGGAAGAGATGCCAGACTAGCTGGAGTCG 970
Db 926 TCATTTATGTCACAAATAGCTAGCCAAATGGAAGAGATGCCAGACTAGCTGGAGTCG 985
QY 971 CTACCTTTGTTCTCCAAAGAGAAATTTGATAGATATTCTGGCTATTGGTGGTCCAAAG 1030
Db 986 CTACCTTTGTTCTCCAAAGAGAAATTTGATAGATATTCTGGCTATTGGTGGTCCAAAG 1045
QY 1031 CTGAAACAACTCCAGTGGTGGTAAATTTAGAAATCTATATGAGAAATGATGAAT 1090
Db 1046 CTGAAACAACTCCAGTGGTGGTAAATTTAGAAATCTATATGAGAAATGATGAAT 1105
QY 1091 CTCAGGTGGAAATTTATGTTATCATCCCTATGTTGGAAAAGAGGCGCAGATTCTAT 1150
Db 1106 CTCAGGTGGAAATTTATGTTATCATCCCTATGTTGGAAAAGAGGCGCAGATTCTAT 1165
QY 1151 TCCGTTATCTTAAACAGGTACAGCAATCTTAAAGTCACTTTTAAAGATGTCAGAAATAA 1210
Db 1166 TCCGTTATCTTAAACAGGTACAGCAATCTTAAAGTCACTTTTAAAGATGTCAGAAATAA 1225
QY 1211 TGATTGATGCTGAGAGAGATCATAGATGTCATAGATAAGGAATTAATTTCAACCTTTTG 1270
Db 1226 TGATTGATGCTGAGAGAGATCATAGATGTCATAGATAAGGAATTAATTTCAACCTTTTG 1285
QY 1271 AGATTCTATTGAAAGAGTTGAATATATTGCCAGAGCTGGATGGACTCCTTGAGGGAAT 1330
Db 1286 AGATTCTATTGAAAGAGTTGAATATATTGCCAGAGCTGGATGGACTCCTTGAGGGAAT 1345
QY 1331 ATGCTTGGTCCATCTCTACTAGATCGCTCCAGACTCGCCCTACAGATAGTGTGATCTCAC 1390
Db 1346 ATGCTTGGTCCATCTCTACTAGATCGCTCCAGACTCGCCCTACAGATAGTGTGATCTCAC 1405
QY 1391 CTGAATATTATTTATCCCAAGTAGAAGATGATGTTATGGAAAGGCGAGAGACTCAITTAGTCTAG 1450

Db 1406 CTGAATATTATTTATCCCAAGTAGAAGATGATTTATGGAAAGGCGAGAGCTCATTTAGTCTAG 1465
QY 1451 TGCCCTGATCTCTGTGAGCGCACTAAATTTATCTATGAAAGAAACACAGACATCTGGATAAATA 1510
Db 1466 TGCCCTGATCTCTGTGAGCGCACTAAATTTATCTATGAAAGAAACACAGACATCTGGATAAATA 1525
QY 1511 TCCATGACATCTTTTCATGTTTTCCTCCAAAGTCAACGAGAGAAATTTGAGTTTATTTTTG 1570
Db 1526 TCCATGACATCTTTTCATGTTTTCCTCCAAAGTCAACGAGAGAAATTTGAGTTTATTTTTG 1585
QY 1571 CCTCTGAATGCAAAACAGGTTTCCGTCATTTATACAAAATTTACATCTATTTTAAAGGAAA 1630
Db 1586 CCTCTGAATGCAAAACAGGTTTCCGTCATTTATACAAAATTTACATCTATTTTAAAGGAAA 1645
QY 1631 GCAAAATATAAACATCCAGTGGTGGCTGCTGCTCCAAGTGATTTCAAGTGTCTCTATCA 1690
Db 1646 GCAAAATATAAACATCCAGTGGTGGCTGCTGCTCCAAGTGATTTCAAGTGTCTCTATCA 1705
QY 1691 AAGAGGAGATAGCAATTTACAGTGGTGGATGGGAAAGTTCTTGGCCCGGCATGGATCTAATA 1750
Db 1706 AAGAGGAGATAGCAATTTACAGTGGTGGATGGGAAAGTTCTTGGCCCGGCATGGATCTAATA 1765
QY 1751 TCCAAGTTGATGAAGTCAGAAAGGCTGGTATATTTTAAAGGCAACCAAGACTCCCTTTTAG 1810
Db 1766 TCCAAGTTGATGAAGTCAGAAAGGCTGGTATATTTTAAAGGCAACCAAGACTCCCTTTTAG 1825
QY 1811 AGCATCACCTGTACGTAGTCAGTTACGTAAATCTTGGAGAGGTGACAAAGGCTGACTGACC 1870
Db 1826 AGCATCACCTGTACGTAGTCAGTTACGTAAATCTTGGAGAGGTGACAAAGGCTGACTGACC 1885
QY 1871 GTGGCTACTCACATCTCTGCTGCATCAGTCAGCACTGTGACTTCTTTTATAGTAAGTATA 1930
Db 1886 GTGGCTACTCACATCTCTGCTGCATCAGTCAGCACTGTGACTTCTTTTATAGTAAGTATA 1945
QY 1931 GTAACAGAGAGATCCACACTGTGTGCTCCCTTTTACAGCTATCAAGTCTGGAAGTACAC 1990
Db 1946 GTAACAGAGAGATCCACACTGTGTGCTCCCTTTTACAGCTATCAAGTCTGGAAGTACAC 2005
QY 1991 CAATTTGCAAAACAAAGGAATTTTGGGCAACATTTTGGATTTCAGCAGGCTCTCTCCCTG 2050
Db 2006 CAATTTGCAAAACAAAGGAATTTTGGGCAACATTTTGGATTTCAGCAGGCTCTCTCCCTG 2065
QY 2051 ACTATACTCTCCAGAAATTTTCTCTTTTGAAGTACTATCTGGATTTTACATGATGGGA 2110
Db 2066 ACTATACTCTCCAGAAATTTTCTCTTTTGAAGTACTATCTGGATTTTACATGATGGGA 2125
QY 2111 TGCTCTAACAGCTCATGATCTACAGCTGGAAGAAATATCTTACTGTCTGTCTCATAT 2170
Db 2126 TGCTCTAACAGCTCATGATCTACAGCTGGAAGAAATATCTTACTGTCTGTCTCATAT 2185
QY 2171 ATGGTGGTCTCAGGTGCAATTAATCGGTTTAAAGGAGTCAAGTATTTCCGCT 2230
Db 2186 ATGGTGGTCTCAGGTGCAATTAATCGGTTTAAAGGAGTCAAGTATTTCCGCT 2245
QY 2231 TGAATACCCCTAGCCCTCTCTAGGTTATGTTGGTGTAGTGATAGACAAAGGGGATCTCTGC 2290
Db 2246 TGAATACCCCTAGCCCTCTCTAGGTTATGTTGGTGTAGTGATAGACAAAGGGGATCTCTGC 2305
QY 2291 ACCGAGGGCTTAAATTTTCAAGCGCTTTAAATATAAAATGGGTCAAATAGAAATTTGACG 2350
Db 2306 ACCGAGGGCTTAAATTTTCAAGCGCTTTAAATATAAAATGGGTCAAATAGAAATTTGACG 2365
QY 2351 ATCAGGTGGAAGGACTCCAAATATCTAGCTTCTCGATATGATTTCAATGACTTAGATCGTG 2410
Db 2366 ATCAGGTGGAAGGACTCCAAATATCTAGCTTCTCGATATGATTTCAATGACTTAGATCGTG 2425
QY 2411 TGGGCACTCCAGCGCTGGTCTTATGAGGATACCTCTCCCTGATGGCAATTAATCAGAGGT 2470
Db 2426 TGGGCACTCCAGCGCTGGTCTTATGAGGATACCTCTCCCTGATGGCAATTAATCAGAGGT 2485
QY 2471 CAGATATCTTCAGGTTGCTATTTGCTGGGCCCCAGTCACTCTGTGGATCTTCTATGATA 2530

Db 2486 CAGATATCTTCAGGGTTGCTATTGCTGGGGCCCGAGTCACTCTGTGGATCTTCTATGATA 2545
Qy 2531 CAGGATACCGGAACGTTATATGGGTCACTGACCGAGATGAACAGGGCTATTACTTAG 2590
Db 2546 CAGGATACCGGAACGTTATATGGGTCACTGACCGAGATGAACAGGGCTATTACTTAG 2605
Qy 2591 GATCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTGACCAAAATCGTTTACTCTCTTAC 2650
Db 2606 GATCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTGACCAAAATCGTTTACTCTCTTAC 2665
Qy 2651 ATGTTTCTCGGATGAGATGTCATTTTGCACATACCAAGTATATTAAGTATTTTATAG 2710
Db 2666 ATGTTTCTCGGATGAGATGTCATTTTGCACATACCAAGTATATTAAGTATTTTATAG 2725
Qy 2711 TGAGGGCTGGAAGCCATATGATTTACAGATCTATCTCTCAGGAGACACAGCATAGAG 2770
Db 2726 TGAGGGCTGGAAGCCATATGATTTACAGATCTATCTCTCAGGAGACACAGCATAGAG 2785
Qy 2771 TTCTGTAATCGGAGAACATTAAGAACTGCACTCTTTTGCATCTCTTCAAGAAAACCTTG 2830
Db 2786 TTCTGTAATCGGAGAACATTAAGAACTGCACTCTTTTGCATCTCTTCAAGAAAACCTTG 2845
Qy 2831 GATCAGTATTTGCTCTTAAAGTGATATAATTTTGACCTGTGTAGAACTCTCTGTAT 2890
Db 2846 GATCAGTATTTGCTCTTAAAGTGATATAATTTTGACCTGTGTAGAACTCTCTGTAT 2905
Qy 2891 ACATGCTATTTAAACCAATGAGGAGTTTAAATCAACAGAAAACACAGAAATGATCATC 2950
Db 2906 ACATGCTATTTAAACCAATGAGGAGTTTAAATCAACAGAAAACACAGAAATGATCATC 2965
Qy 2951 ACATTTTGATCTGCGATCTTAACATCTACTCTCTGAAAATAAATTTGTGTCATGACGG 3010
Db 2966 ACATTTTGATCTGCGATCTTAACATCTACTCTCTGAAAATAAATTTGTGTCATGACGG 3025
Qy 3011 GTCTACCGTTTGTGTAGTAATCTAATACCTTTAACCACCATGCTCAAAATCAAAATGATA 3070
Db 3026 GTCTACCGTTTGTGTAGTAATCTAATACCTTTAACCACCATGCTCAAAATCAAAATGATA 3085
Qy 3071 CATATCTCGAGAGACCCAGCAATACCAAGAAATTAAGTATTAAGTATTAAGTATTAAG 3120
Db 3086 CATATCTCGAGAGACCCAGCAATACCAAGAAATTAAGTATTAAGTATTAAGTATTAAG 3135

RESULT 5
ID ABK83327 standard; cDNA; 4829 BP.
XX AC ABK83327;
XX DT 12-AUG-2002 (first entry)
XX DE cDNA encoding human DPRP-1 splice variant #3.
XX KW Human; serine protease; dipeptidyl peptidase IV-related protein; DPRP;
KW DPPIV; infection; human immunodeficiency virus; HIV-1; HIV-2; pain;
KW diabetes; infertility; obesity; anorexia; Parkinson's disease; stroke;
KW heart failure; hypertension; urinary retention; osteoporosis; cancer;
KW ulcer; allergy; cancer; psychotic disorder; neurological disorder;
KW dyskinnesia; reproductive disorder; inflammatory disorder;
KW metabolic disorder; gene; ss.
XX OS Homo sapiens.
XX PN WO200231134-A2.
XX PD 18-APR-2002.
XX PF 12-OCT-2001; 2001WO-US031874.
XX PR 12-OCT-2000; 2000US-0240117P.
XX PA (FERR) FERRING BV.
XX

PI Qi S, Akinsanya KO, Riviere PJ, Junien J;
XX WPI; 2002-444178/47.
DR P-PSDB; ABG61596.
XX
PT New dipeptidyl peptidase IV-related proteins and nucleic acids encoding
the proteins, useful for treating e.g. fungal, bacterial, protozoan and
viral infections, cancers, allergies, neurological disorders, or pain.
XX Disclosure; Page 65-66; 113pp; English.
XX
CC The present invention relates to the isolation of novel human serine
proteases referred to as dipeptidyl peptidase IV (DPPIV)-related proteins
(DPRP). The dipeptidyl peptidase IV-related proteins (DPRP) and nucleic
acids encoding them are useful for treating infections such as fungal,
bacterial, protozoan and viral infections, particularly infections caused
by human immunodeficiency virus (HIV-1 or HIV-2), pain, diabetes,
precocious puberty, infertility, obesity, anorexia, bulimia, Parkinson's
disease, acute heart failure, hypotension, hypertension, urinary
retention, osteoporosis, angina pectoris, stroke, ulcers, asthma,
allergies, cancers, migraine, vomiting, psychotic and neurological
disorders (e.g. anxiety, dementia, or schizophrenia), and dyskinesias.
CC These may also be used in discovering therapeutic agents for the
treatment of reproductive, inflammatory and metabolic disorders. ABK83322
CC -ABK83343 encode human DPRP proteins
XX
SQ Sequence 4829 BP; 1466 A; 886 C; 1017 G; 1460 T; 0 U; 0 Other;
Query Match 92.6%; Score 2889; DB 6; Length 4829;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3119; Conservative 0; Mismatches 1; Indels 2; Gaps 1;
Qy 1 AAGTGTCTAAAGCCTCCGAGGCCAAGCCGCTGCTACTCCCGCGCTGTTCTTAGTGGCG 60
Db 1 AAGTGTCTAAAGCCTCCGAGGCCAAGCCGCTGCTACTCCCGCGCTGTTCTTAGTGGCG 60
Qy 61 CGTTTCGCGCGCTGGTGTGTACCGCGCGCGCGCGAGAGCCACTGCACACGAGGACCG 120
Db 61 CGTTTCGCGCGCTGGTGTGTACCGCGCGCGCGCGAGAGCCACTGCACACGAGGACCG 120
Qy 121 GAGTGGAGCGCGCGAGCATGAAGCGCGCGCGCGCTCCATAGCGCACCTCGGACGG 180
Db 121 GAGTGGAGCGCGCGAGCATGAAGCGCGCGCGCTCCATAGCGCACCTCGGACGG 180
Qy 181 TCCGGCGCGCGCGCGAGGAAATGCAACATGGCAGCAGCAATGGAACAGAACAG 240
Db 181 TCCGGCGCGCGCGCGAGGAAATGCAACATGGCAGCAGCAATGGAACAGAACAG 240
Qy 241 CTGGGTCTTGAGATATTTGAAACTGCGGA CTGTGAGGAGAAATTTGAATCAAGATCGG 300
Db 241 CTGGGTCTTGAGATATTTGAAACTGCGGA CTGTGAGGAGAAATTTGAATCAAGATCGG 300
Qy 301 CCTAAATGGAGCCTTTTATGTTGAGCGGTATCTTGGAGTCAGCTTAAAGTCTT 360
Db 301 CCTAAATGGAGCCTTTTATGTTGAGCGGTATCTTGGAGTCAGCTTAAAGTCTT 360
Qy 361 GCCGATACCGAAAAATATCATGCTACATGCTTAAAGCAGCAGCATGTTTATGTTT 420
Db 361 GCCGATACCGAAAAATATCATGCTACATGCTTAAAGCAGCAGCATGTTTATGTTT 420
Qy 421 GTGAGGAGGATGATCCAGATGACCTCTTTCAGACAGAAATCTATTCTTCCCATGTCT 480
Db 421 GTGAGGAGGATGATCCAGATGACCTCTTTCAGACAGAAATCTATTCTTCCCATGTCT 480
Qy 481 GTGAGAACAGAGAAAAATACACTGTTTATTTCTGAAATTCGCAAAATATCAATAGACA 540
Db 481 GTGAGAACAGAGAAAAATACACTGTTTATTTCTGAAATTCGCAAAATATCAATAGACA 540
Qy 541 GCAGTCTTAATGCTCTCTTGGAAAGCCTCTTTTGGATCTTTTTCAGGCAACACTGACTAT 600
Db 541 GCAGTCTTAATGCTCTCTTGGAAAGCCTCTTTTGGATCTTTTTCAGGCAACACTGACTAT 600
Qy 601 GGAATGTATTCTCGAGAGAGAAAGAACTATTAAAGAGAAAGAAAGCACTTGGACAGTCGGA 660

Db 601 GGAATGTATTCTCAGAGAAGAACTATTAAAGAGAAAGAAAACGCAATGGAAACAGTCGGA 660
Qy 661 ATTGCTCTTACGATTATCACCAGGAGTGGAAACATTTCTGTTTCAAGCCGGTAGTGGGA 720
Db 661 ATTGCTCTTACGATTATCACCAGGAGTGGAAACATTTCTGTTTCAAGCCGGTAGTGGGA 720
Qy 721 ATTTATCAGCTAAAGATGAGGCGCACAAAGGATTTACGCCAACACCTTTTAAGGCCCAAT 780
Db 721 ATTTATCAGCTAAAGATGAGGCGCACAAAGGATTTACGCCAACACCTTTTAAGGCCCAAT 780
Qy 781 CTAGTGGAACTAGTTGTCGCAACATACGGAATGGATTCAAAAATTAATGCCCGCTGATCCA 840
Db 781 CTAGTGGAACTAGTTGTCGCAACATACGGAATGGATTCAAAAATTAATGCCCGCTGATCCA 840
Qy 841 GACTGGATTTGCTTTTATACATAGCAACGATATTTGGATATCTAACCATCGTAACAGAGAA 900
Db 841 GACTGGATTTGCTTTTATACATAGCAACGATATTTGGATATCTAACCATCGTAACAGAGAA 900
Qy 901 GAAAGGAGACTCACTTATGTGCAAAATGAGCTAGCCAAACATGGAAGAAGATGCCAGATCA 960
Db 901 GAAAGGAGACTCACTTATGTGCAAAATGAGCTAGCCAAACATGGAAGAAGATGCCAGATCA 960
Qy 961 GCTGGAGTCTGCTACCTTTGTTCTCCAAAGAGAAATTTGATAGATATTTCTGGCTATTGGTGG 1020
Db 961 GCTGGAGTCTGCTACCTTTGTTCTCCAAAGAGAAATTTGATAGATATTTCTGGCTATTGGTGG 1020
Qy 1021 TGTCCAAAAGCTGAAACAACTCCAGTGGTGTAAATTTCTTAGAATTTCTATATGAGAA 1080
Db 1021 TGTCCAAAAGCTGAAACAACTCCAGTGGTGTAAATTTCTTAGAATTTCTATATGAGAA 1080
Qy 1081 AATGATCAATCTCAGGTGGAAATTTATCATGTTTACATCCCTATGTTGGAAACAGGAGG 1140
Db 1081 AATGATCAATCTCAGGTGGAAATTTATCATGTTTACATCCCTATGTTGGAAACAGGAGG 1140
Qy 1141 GCAGATTCATCTCGTTATCTTAAACAGGTACAGCAATCTTAAAGTCACTTTTAAGATG 1200
Db 1141 GCAGATTCATCTCGTTATCTTAAACAGGTACAGCAATCTTAAAGTCACTTTTAAGATG 1200
Qy 1201 TCAGAAATATGATGATCTGAAAGAGATCATAGATGTCATAGATAAGAACTAAAT 1260
Db 1201 TCAGAAATATGATGATCTGAAAGAGATCATAGATGTCATAGATAAGAACTAAAT 1260
Qy 1261 CAACCTTTTGAGATCTTATTTGAAGAGTGAATATATGTCAGAGCTGGATGACTCT 1320
Db 1261 CAACCTTTTGAGATCTTATTTGAAGAGTGAATATATGTCAGAGCTGGATGACTCT 1320
Qy 1321 GAGGAAATATGCTTGGTCCATCTACTAGATCGCTCCAGACTCGCTTACAGATAGTG 1380
Db 1321 GAGGAAATATGCTTGGTCCATCTACTAGATCGCTCCAGACTCGCTTACAGATAGTG 1380
Qy 1381 TTGATCTCACTGAATTTATTTATCCAGTAGAAGATGATGTTATGAAAAGCGAGACTC 1440
Db 1381 TTGATCTCACTGAATTTATTTATCCAGTAGAAGATGATGTTATGAAAAGCGAGACTC 1440
Qy 1441 ATTGAGTCAGTGCCTGATCTGTGACGCCACTAATTTATCTATGAAGAAACACAGACATC 1500
Db 1441 ATTGAGTCAGTGCCTGATCTGTGACGCCACTAATTTATCTATGAAGAAACACAGACATC 1500
Qy 1501 TGGATAAATATCATGACATCTTTTCATGTTTTCCCAAAGTCACGAAGGAAATTTGAG 1560
Db 1501 TGGATAAATATCATGACATCTTTTCATGTTTTCCCAAAGTCACGAAGGAAATTTGAG 1560
Qy 1561 TTTATTTTGGCTCTGAAATGAAAACAGGTTTCGGTCAATTTATACAAAATTTACATCTAT 1620
Db 1561 TTTATTTTGGCTCTGAAATGAAAACAGGTTTCGGTCAATTTATACAAAATTTACATCTAT 1620
Qy 1621 TTTAAAGGAAACAAATATAACGATCCAGTGGTGGCTGCTCTCAAGATTTCAAG 1680
Db 1621 TTTAAAGGAAACAAATATAACGATCCAGTGGTGGCTGCTCTCAAGATTTCAAG 1680
Qy 1681 TGTCTCTACAAAGAGATAGCAATTTACCAGTGGTGAATGGGAAGTTCTTGGCCGGCAT 1740

Db 1681 TGTCTCTACAAAGAGGAGATAGCAATTTACAGTGGTGAATGGGAAGTTCTTGGCCGGCAT 1740
Qy 1741 GGATCTAATATCCAAGTTGATGAAGTCAGAAAGCGCTGGTATATTTTGAAGGACCAACAAAGAC 1800
Db 1741 GGATCTAATATCCAAGTTGATGAAGTCAGAAAGCGCTGGTATATTTTGAAGGACCAACAAAGAC 1800
Qy 1801 TCCCTCTTTAGAGCATCACCTGTACGTAGTCAGTTACGTAAATCTCTGGAGAGGTGACAAGG 1860
Db 1801 TCCCTCTTTAGAGCATCACCTGTACGTAGTCAGTTACGTAAATCTCTGGAGAGGTGACAAGG 1860
Qy 1861 CTGACTCAACCGTGGCTACTCACAATTTCTGCTCATCAGTCAGCAGCTGACTTCTTTATA 1920
Db 1861 CTGACTCAACCGTGGCTACTCACAATTTCTGCTCATCAGTCAGCAGCTGACTTCTTTATA 1920
Qy 1921 AGTAAGTATAGTAACAGAGAAATCCACACTGCTGTGTCCTTTTGAAGTACTACTGATTCCT 1980
Db 1921 AGTAAGTATAGTAACAGAGAAATCCACACTGCTGTGTCCTTTTGAAGTACTACTGATTCCT 1980
Qy 1981 GNAGATGACCCAACTTTGCAAAAACAAGGAATTTTGGGCCACATTTTGGATTGACGAGGT 2040
Db 1981 GNAGATGACCCAACTTTGCAAAAACAAGGAATTTTGGGCCACATTTTGGATTGACGAGGT 2040
Qy 2041 CCTCTTCTGACTACTCTCTCGAANAATTTCTCTTTTGAAGTACTACTGATTCCTGATTCACA 2100
Db 2041 CCTCTTCTGACTACTCTCTCGAANAATTTCTCTTTTGAAGTACTACTGATTCCTGATTCACA 2100
Qy 2101 TTGTATGGGATGCTCTAAGGCTCATGATCTAAGCCTGGAAAAGAAATATCTCTACTGTG 2160
Db 2101 TTGTATGGGATGCTCTAAGGCTCATGATCTAAGCCTGGAAAAGAAATATCTCTACTGTG 2160
Qy 2161 CTGTTCATATATGTTGG--TCCTCAGGTGCAAGTTGGTGAATAATCGGTTTAAAGGAGTCA 2218
Db 2161 CTGTTCATATATGTTGGTCTCTCTCAGGTGCAAGTTGGTGAATAATCGGTTTAAAGGAGTCA 2220
Qy 2219 AGTAATTCGCTTGAATACCTTAGCCTCTCTAGGTTATGTTGTTAGTATGATGACAACA 2278
Db 2221 AGTAATTCGCTTGAATACCTTAGCCTCTCTAGGTTATGTTGTTAGTATGATGACAACA 2280
Qy 2279 GGGGATCTCTGTCAACGAGGCTTAAATTTGAAGGCGCTTTTAAATATAAAATGGTCAAA 2338
Db 2281 GGGGATCTCTGTCAACGAGGCTTAAATTTGAAGGCGCTTTTAAATATAAAATGGTCAAA 2340
Qy 2339 TAGAAATTCAGCATCAGGTGGAAGGACTCCAAATATCTAGCTTCTCGATATGATTTCATTG 2398
Db 2341 TAGAAATTCAGCATCAGGTGGAAGGACTCCAAATATCTAGCTTCTCGATATGATTTCATTG 2400
Qy 2399 ACTTAGATCGTGTGGGCATCCACGCTGCTCTATGAGGATACCTCTCCCTGATGGCAT 2458
Db 2401 ACTTAGATCGTGTGGGCATCCACGCTGCTCTATGAGGATACCTCTCCCTGATGGCAT 2460
Qy 2459 TAATGAGAGGTCAGATATCTTCAGGTTGCTATTTGCTGGGCCCCAGTCACCTCTGTGGA 2518
Db 2461 TAATGAGAGGTCAGATATCTTCAGGTTGCTATTTGCTGGGCCCCAGTCACCTCTGTGGA 2520
Qy 2519 TCTTCTATGATACAGGATACAGGAACTTATATGGGTCACTCTGACCCAGCAGAAATGAACAGG 2578
Db 2521 TCTTCTATGATACAGGATACAGGAACTTATATGGGTCACTCTGACCCAGCAGAAATGAACAGG 2580
Qy 2579 GCTATTTAGTAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTGAAACCAAAATCGTT 2638
Db 2581 GCTATTTAGTAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTGAAACCAAAATCGTT 2640
Qy 2639 TACTGCTCTTACATGTTTCTCGATGAGATGTCATTTTGAACATACAGTATATTAC 2698
Db 2641 TACTGCTCTTACATGTTTCTCGATGAGATGTCATTTTGAACATACAGTATATTAC 2700
Qy 2699 TGAGTTTTTATGAGGCTGGAAGCCATATGATTTTACAGATCTATCTCTCAGGAGAGAC 2758
Db 2701 TGAGTTTTTATGAGGCTGGAAGCCATATGATTTTACAGATCTATCTCTCAGGAGAGAC 2760
Qy 2759 ACAGATAAGAGTTCTCTGAATCGGAGAACATTTATGAATGCAATCTTTTGGCACTACCTTC 2818
Db 2761 ACAGATAAGAGTTCTCTGAATCGGAGAACATTTATGAATGCAATCTTTTGGCACTACCTTC 2820

QY 2819 AAGAAACCTTGGATCAGCTATTGCTGCTCTAAAGTGATATAATTTTGACCTGTGTAGA 2878
DB 2821 AAGAAACCTTGGATCAGCTATTGCTGCTCTAAAGTGATATAATTTTGACCTGTGTAGA 2880
QY 2879 ACTCTCTGGTATACACTGGCTATTTAAACCAATGAGGAGTTTAAATCAACAGAAACACA 2938
DB 2881 ACTCTCTGGTATACACTGGCTATTTAAACCAATGAGGAGTTTAAATCAACAGAAACACA 2940
QY 2939 GAAATTGATCATCATATTTTGATACCTGCGCATGTAAACATCTACTCTCTGAAATAAATGTGG 2998
DB 2941 GAAATTGATCATCATATTTTGATACCTGCGCATGTAAACATCTACTCTCTGAAATAAATGTGG 3000
QY 2999 TGCCATGAGGGGTCTACGGTTTGTGGTAGTAACTTAATACCTTAACCCCATGCTCAA 3058
DB 3001 TGCCATGAGGGGTCTACGGTTTGTGGTAGTAACTTAATACCTTAACCCCATGCTCAA 3060
QY 3059 AATCAATGATACATATTTCTGAGAGACCCAGCAATACCATAGAAATTAACAAAAAAA 3118
DB 3061 AATCAATGATACATATTTCTGAGAGACCCAGCAATACCATAGAAATTAACAAAAAAA 3120
QY 3119 AA 3120
DB 3121 AA 3122

RESULT 6
ADV43981
ID ADV43981 standard; cDNA; 2649 BP.
AC ADV43981;
XX
DT 10-MAR-2005 (first entry)
XX
DE Human psychoneuroendocrinimmune expressed sequence tag SEQ ID NO 1609.
XX
KW microarray; psychoneuroendocrinimmune; chronic fatigue;
KW non-insulin dependent diabetes; allergy; immune disorder; inflammation;
KW cancer; neoplasm; infection; expressed sequence tag; ss.
XX
OS Homo sapiens.
XX
XX WO2004108899-A2.
XX
PD 16-DEC-2004.
XX
XX 04-JUN-2004; 2004WO-US017686.
XX
XX 04-JUN-2003; 2003US-0475915P.
XX
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Nicholson A, Vernon SD;
XX
XX WPI; 2005-031682/03.
XX
XX New microarray comprising probes for genes involved in
PT psychoneuroendocrinimmune (PNI) activity, useful in diagnosing a
PT condition associated with PNI activity, e.g., inflammatory or infectious
PT diseases.
XX
XX Claim 1; SEQ ID NO 1609; 254pp; English.
XX
XX The invention relates to a new microarray which comprises probes for
CC genes involved in psychoneuroendocrinimmune (PNI) activity. The
CC microarray is useful in diagnosing a condition associated with PNI
CC activity, such as CFS, type-2 diabetes, allergic condition, inflammation,
CC cancer and infection. The present sequence represents a
CC psychoneuroendocrinimmune gene expressed sequence tag. Note the
CC specification mentions SEQ ID NO of up to 3314 but only sequences up to
CC SEQ ID NO 1829 are provided.
XX
XX Sequence 2649 BP; 804 A; 514 C; 585 G; 746 T; 0 U; 0 Other;

Query Match 83.3%; Score 2598; DB 14; Length 2649;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2648; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 214 ATGCAGCAGCAATGGAAACAGAAACAGAGCTGGGTGTGGAGATATTTGAAATCGCGACTGT 273
DB 1 ATGCAGCAGCAATGGAAACAGAAACAGAGCTGGGTGTGGAGATATTTGAAATCGCGACTGT 60
QY 274 GAGGAGAATATTTGAATCAGAGGATCGGCCTAAATTTGGAGCCCTTTTATTTGAGGGGTAT 333
DB 61 GAGGAGAATATTTGAATCAGAGGATCGGCCTAAATTTGGAGCCCTTTTATTTGAGGGGTAT 120
QY 334 TCCTGGAGTCAGCTTAAAGCTGCTTCGGATACACAGAAATATCATCGGCTACATGATG 393
DB 121 TCCTGGAGTCAGCTTAAAGCTGCTTCGGATACACAGAAATATCATCGGCTACATGATG 180
QY 394 GCTAAGGCACCATATGATTTTCATGTTGTGAAGAGAAATGATCCAGATGGACCTCATTTCA 453
DB 191 GCTAAGGCACCATATGATTTTCATGTTGTGAAGAGAAATGATCCAGATGGACCTCATTTCA 240
QY 454 GACAGAATCTATTACCTTGCCATGCTGCTGAGAAACAGAGAAATATCATGTTTATTTCT 513
DB 241 GACAGAATCTATTACCTTGCCATGCTGCTGAGAAACAGAGAAATATCATGTTTATTTCT 300
QY 514 GAAATTCCTCAAACTATCAATAGAGCAGCTTAAATGCTCTCTTGGAGCCCTTTTGTG 573
DB 301 GAAATTCCTCAAACTATCAATAGAGCAGCTTAAATGCTCTCTTGGAGCCCTTTTGTG 360
QY 574 GATCTTTTTCAGCAACACTGGAATGATGTAATGTTCTCGAGAGAGAACTATTAAGA 633
DB 361 GATCTTTTTCAGCAACACTGGAATGATGTAATGTTCTCGAGAGAGAACTATTAAGA 420
QY 634 GAAAGAAAACGCAATGGAAACAGTCGGAATTTGCTTTCAGATTTATCACCAAGAAAGTGA 693
DB 421 GAAAGAAAACGCAATGGAAACAGTCGGAATTTGCTTTCAGATTTATCACCAAGAAAGTGA 480
QY 694 ACATTTCTGTTTCAAGCCGGTAGTGGAAATTTATCAGTAAAGAGTGGAGGGCCACAAGGA 753
DB 481 ACATTTCTGTTTCAAGCCGGTAGTGGAAATTTATCAGTAAAGAGTGGAGGGCCACAAGGA 540
QY 754 TTTACGCAACAACTTTAAGCCCAATCTAGTGGAAACCTAGTGTGCCAAATACAGGATG 813
DB 541 TTTACGCAACAACTTTAAGCCCAATCTAGTGGAAACCTAGTGTGCCAAATACAGGATG 600
QY 814 GATCCAAAATTTATGCCCCGCTGATCCAGACTGGATGCTTTTATACATAGCAACCATATT 873
DB 601 GATCCAAAATTTATGCCCCGCTGATCCAGACTGGATGCTTTTATACATAGCAACCATATT 660
QY 874 TGGATATCTAACATCGTAACACAGAGAAAGAGAGACTCACTTATGTGCACAATGAGCTA 933
DB 661 TGGATATCTAACATCGTAACACAGAGAAAGAGAGACTCACTTATGTGCACAATGAGCTA 720
QY 934 GCCAACATGGAAGAAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAAGAAGAA 993
DB 721 GCCAACATGGAAGAAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAAGAAGAA 780
QY 994 TTTGATAGATATTTCTGGCTATTTGGTGTGTCCTCAAAAGCTGAAACAACTCCAGTGGTGT 1053
DB 781 TTTGATAGATATTTCTGGCTATTTGGTGTGTCCTCAAAAGCTGAAACAACTCCAGTGGTGT 840
QY 1054 AAAATTTTGAATTTCTATATGAAGAAATGATGAATCTGAGGTGGAAATTTATCATGTT 1113
DB 841 AAAATTTTGAATTTCTATATGAAGAAATGATGAATCTGAGGTGGAAATTTATCATGTT 900
QY 1114 ACATCCCCCTATTTGGAAACAGAGGGCGAGATTCCTCGGTATCTCTAAACAGGTACA 1173
DB 901 ACATCCCCCTATTTGGAAACAGAGGGCGAGATTCCTCGGTATCTCTAAACAGGTACA 960
QY 1174 GCATAATCTAAAGTCACCTTTTAAAGATGTCAGAAATTAATGATGATCTCAAGGAAGATC 1233
DB 961 GCATAATCTAAAGTCACCTTTTAAAGATGTCAGAAATTAATGATGATCTCAAGGAAGATC 1020

QY 1234 ATAGATGTCATAGATAAGGAACAAATTCACACCTTTTGGAGATTCCTATTGTAAGGAGTTGAA 1293
Db 1021 ATAGATGTCATAGATAAGGAACAAATTCACACCTTTTGGAGATTCCTATTGTAAGGAGTTGAA 1080
QY 1294 TATATTGCGACAGCTGGATGGACTCCTGAGGGAATAATATGCTGGTCCATCTACTAGAT 1353
Db 1081 TATATTGCGACAGCTGGATGGACTCCTGAGGGAATAATATGCTGGTCCATCTACTAGAT 1140
QY 1354 CGCTCCAGAGTCCGCTTACAGATAGTGTGATCTCACCTGAAATTAATATCCCAAGTAGAA 1413
Db 1141 CGCTCCAGAGTCCGCTTACAGATAGTGTGATCTCACCTGAAATTAATATCCCAAGTAGAA 1200
QY 1414 GATGATGTTATGGAAGGCGAGAGACTCATTTGAGTCAGTCCCTGATCTCTGTGACGCCACTA 1473
Db 1201 GATGATGTTATGGAAGGCGAGAGACTCATTTGAGTCAGTCCCTGATCTCTGTGACGCCACTA 1260
QY 1474 ATTTATCTATGAAGAAACAAACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTTTT 1533
Db 1261 ATTTATCTATGAAGAAACAAACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTTTT 1320
QY 1534 CCCCAAGTCACGAAGAGGAAATTTGAGTTTATTTTTTGGCTCTGAAATGCAAAACAGGTTTC 1593
Db 1321 CCCCAAGTCACGAAGAGGAAATTTGAGTTTATTTTTTGGCTCTGAAATGCAAAACAGGTTTC 1380
QY 1594 CGTCATTTATCAAAATTTACATCTATTTTAAAGGAAGCAATATATAACGATCCAGTGGT 1653
Db 1381 CGTCATTTATCAAAATTTACATCTATTTTAAAGGAAGCAATATATAACGATCCAGTGGT 1440
QY 1654 GGGCTGCTCTGCCAAGTGATTTCAAGTCTCCTATCAAGAGGAGATAGCAATTTACCAGT 1713
Db 1441 GGGCTGCTCTGCCAAGTGATTTCAAGTCTCCTATCAAGAGGAGATAGCAATTTACCAGT 1500
QY 1714 GGTCAATGGGAAGTCTTGCGCGGATGATCTATATATCAAGTTGATGAAGTCAGAAGG 1773
Db 1501 GGTCAATGGGAAGTCTTGCGCGGATGATCTATATCAAGTTGATGAAGTCAGAAGG 1560
QY 1774 CTGCTATATTTGGAAGCCAAAGACTCCCTTTTAGAGCATCACTGTACGTAGTCAGT 1833
Db 1561 CTGCTATATTTGGAAGCCAAAGACTCCCTTTTAGAGCATCACTGTACGTAGTCAGT 1620
QY 1834 TACGTAAATCCTGGAGAGTGACAAGGCTGACTGACCGTGGCTACTCACATTTCTTGCTGC 1893
Db 1621 TACGTAAATCCTGGAGAGTGACAAGGCTGACTGACCGTGGCTACTCACATTTCTTGCTGC 1680
QY 1894 ATCAGTCAGCACTGTGACTTCTTTTATAAGTAAGTATAGTAACCAAGAAATCCACACTGT 1953
Db 1681 ATCAGTCAGCACTGTGACTTCTTTTATAAGTAAGTATAGTAACCAAGAAATCCACACTGT 1740
QY 1954 GTGTCCCTTTTACAAGCTATCAAGTCTGGAAGATGACCCAACTTCGAAAACAAAAGGAATTT 2013
Db 1741 GTGTCCCTTTTACAAGCTATCAAGTCTGGAAGATGACCCAACTTCGAAAACAAAAGGAATTT 1800
QY 2014 TGGGCCACCAATTTTGGATTCAGCAGGTCTCTCTCTGACTATCTCTCCAGAAATTTTC 2073
Db 1801 TGGGCCACCAATTTTGGATTCAGCAGGTCTCTCTCTGACTATCTCTCCAGAAATTTTC 1860
QY 2074 TCTTTTGAAGTACTACTGATTTACATGTATGGGATGCTCTCAAGCCTCATGATCTA 2133
Db 1861 TCTTTTGAAGTACTACTGATTTACATGTATGGGATGCTCTCAAGCCTCATGATCTA 1920
QY 2134 CAGCTGGAAAGAAATATCCTACTGTGCTTTTATATATGTTGTTCTCAGGTGCAAGTTG 2193
Db 1921 CAGCTGGAAAGAAATATCCTACTGTGCTTTTATATATGTTGTTCTCAGGTGCAAGTTG 1980
QY 2194 GTGAATAATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCCCTAGCCTCTCTAGGT 2253
Db 1981 GTGAATAATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCCCTAGCCTCTCTAGGT 2040
QY 2254 TATGTGGTTGTAGTAGACAAACAGGGGATCCTGTCAACGAGGGCTTTAAATTTGAAGGC 2313
Db 2041 TATGTGGTTGTAGTAGACAAACAGGGGATCCTGTCAACGAGGGCTTTAAATTTGAAGGC 2100
QY 2314 GCCTTTTAAATATGAGTGGTCAAAATAGAAATTTGACGATCAGGTGGAAGGACTCCAATAT 2373

Db 2101 GCCTTTTAAATATGAGTGGTCAAAATAGAAATTTGACGATCAGGTGGAAGGACTCCAATAT 2160
QY 2374 CTAGCTTCTCGATATGATTTTCATTTAGATCTTAGATCGTGTGGGCATCCAGGCTGGTCCCTAT 2433
Db 2161 CTAGCTTCTCGATATGATTTTCATTTAGATCTTAGATCGTGTGGGCATCCAGGCTGGTCCCTAT 2220
QY 2434 GGAGGATACCTCTCCCTCGATGGCATTAATGACAGAGTCAAGATATCTTCAGGGTTGCTATT 2493
Db 2221 GGAGGATACCTCTCCCTCGATGGCATTAATGACAGAGTCAAGATATCTTCAGGGTTGCTATT 2280
QY 2494 GCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAACGTTTATATG 2553
Db 2281 GCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAACGTTTATATG 2340
QY 2554 GGTCAACCTGACACAGAAATGAACAGGGCTATTACTTAGGATCTGTGGCATGCAAGCAGAA 2613
Db 2341 GGTCAACCTGACACAGAAATGAACAGGGCTATTACTTAGGATCTGTGGCATGCAAGCAGAA 2400
QY 2614 AAGTTTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGTTTCTCGATGAGNATGTC 2673
Db 2401 AAGTTTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGTTTCTCGATGAGNATGTC 2460
QY 2674 CATTTGCGACATACCACTATATTAAGTCTTTTGTAGAGGCTGGAAAGCCATATGAT 2733
Db 2461 CATTTGCGACATACCACTATATTAAGTCTTTTGTAGAGGCTGGAAAGCCATATGAT 2520
QY 2734 TTACAGATCTATCCTCAGAGAGACACAGCATAAAGAGTTCCTGAAATCGGAGAAACATTAT 2793
Db 2521 TTACAGATCTATCCTCAGAGAGACACAGCATAAAGAGTTCCTGAAATCGGAGAAACATTAT 2580
QY 2794 GAACCTGATCTTTTGGTCACTACCTTCAAGAAAAACCTTGGATCACGTTTGTGCTCTAAAA 2853
Db 2581 GAACCTGATCTTTTGGTCACTACCTTCAAGAAAAACCTTGGATCACGTTTGTGCTCTAAAA 2640
QY 2854 GTGATATAA 2862
Db 2641 GTGATATAA 2649
RESULT 7
AAH99935
ID AAH99935 standard; cDNA; 2643 bp.
XX AAH99935;
XX
XX
XX 12-APR-2002 (first entry)
XX Coding sequence of 21953 human prollyl oligopeptidase.
DE 21953 prollyl oligopeptidase; antibody; proline; endopeptidase; cancer;
KW cardiovascular disease; autoimmune disease; atopic allergy;
KW neuronal disorder; vascular disorder; prostate disorder; cytostatic;
KW antidiabetic; antiarthritic; antiasthmatic; antiinflammatory;
KW diabetes mellitus; arthritis; multiple sclerosis; asthma;
KW Grave's disease; neuronal disorder; demyelinating disease; ss.
OS Homo sapiens.
XX
XX WO200179473-A2.
XX
XX 25-OCT-2001.
XX
XX 11-APR-2001; 2001WO-US040483.
XX 18-APR-2000; 2000US-0197508P.
XX
XX (MILL-) MILLENNIUM PHARM INC.
XX Meyers RA, Williamson M;
XX WPI; 2002-034353/04.
XX P-PSDB; AAG78415.
DR

XX New polypeptides 21953, member of human prolyl oligopeptidase family,
PT useful as diagnostic targets and therapeutic agents for controlling
PT cancer, lymphoma and leukemia.
XX
XX
PS Claim 2; Page; 121pp; English.
XX
CC This invention relates to an isolated 21953 human prolyl oligopeptidase.
CC Which is cytostatic, antidiabetic, antiarthritic, neuroprotective,
CC antihydroid, dermatological, antipsoriatic, antiasthmatic,
CC ophthalmological, antiinflammatory, nootropic, antiparkinsonian,
CC anticonvulsant, gynaecological, vasotropic, antianginal, cardiant,
CC antiatherosclerotic, anorectic and metabolic in its action. Uses include
CC gene therapy, expression or activity of 21953 protein modulator, it is
CC useful for identifying a compound which binds to it and can be used in
CC preventing, treating or detecting a cellular proliferative or
CC differentiative disorder. The 21953 molecules can act as novel diagnostic
CC targets and therapeutic agents for controlling disorders associated with
CC the aberrant activity or degradation of peptide hormones e.g., disorders
CC associated with cell differentiation and proliferation such as cancer,
CC immune function, reproductive, neurological and cardiovascular function.
CC The 21953 molecules are thus useful for treating and preventing cellular
CC proliferative and differentiative disorders, haematopoietic neoplastic
CC disorders, immune disorders such as autoimmune diseases, diabetes
CC mellitus, arthritis, multiple sclerosis, asthma, Grave's disease,
CC neuronal disorders, demyelinating diseases, vascular disorders and
CC metabolism or pain disorders. This sequence represents the cDNA encoding
CC sequence of 21953 human prolyl oligopeptidase. This sequence represents
CC the coding sequence of 21953, being the sequence in between the start and
CC the stop codon of the sequence represented in AAH99934
XX
SQ Sequence 2643 BP; 800 A; 514 C; 585 G; 744 T; 0 U; 0 Other;

Query Match 81.4%; Score 2541; DB 6; Length 2643;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2641; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
217 GCAGCAGCAATGAAACAGACAGCTGGGTGTGAGATATTTGAACTGGGCTGTGAG 276
1 GCAGCAGCAATGAAACAGACAGCTGGGTGTGAGATATTTGAACTGGGCTGTGAG 60
277 GAGAAATTTGAATCAGAGATCGGCTAAATTTGAGCGCTTTTATTTGAGCGGTATTC 336
61 GAGAATATTTGAATCAGAGATCGGCTAAATTTGAGCGCTTTTATTTGAGCGGTATTC 120
337 TGGAGTCAGCTTAAAGCTGCTTGCAGATACAGAAATATCATGGCTACATGAGGCT 396
121 TGGAGTCAGCTTAAAGCTGCTTGCAGATACAGAAATATCATGGCTACATGAGGCT 180
397 AAGCACCACATGATTTCAATTTGTTGTAAGAGAAATGATCCAGATGGACCTCATTCAGAC 456
181 AAGCACCACATGATTTCAATTTGTTGTAAGAGAAATGATCCAGATGGACCTCATTCAGAC 240
457 AGAATCTATTACCTTGCATGCTGTGTGAGAACAGAAATACACTGTTTATTTCTGAA 516
241 AGAATCTATTACCTTGCATGCTGTGTGAGAACAGAAATACACTGTTTATTTCTGAA 300
517 ATTCCTCAAAATCATATAGCAGCAGCTTAAATGCTCTCTTGAAGCCCTCTTTTGGAT 576
301 ATTCCTCAAAATCATATAGCAGCAGCTTAAATGCTCTCTTGAAGCCCTCTTTTGGAT 360
577 CTTTTTCAGGCAACACTGGACTATGGAATGATTTCTCGAGAAAGAAACTATTAAAGAA 636
361 CTTTTTCAGGCAACACTGGACTATGGAATGATTTCTCGAGAAAGAAACTATTAAAGAA 420
637 AAAAAACCATTTGAAACAGTCGGAATTTGCTTTACGATTATCAACCAAGGAAGTGAACA 696
421 AGAAAAACCATTTGAAACAGTCGGAATTTGCTTTACGATTATCAACCAAGGAAGTGAACA 480
697 TTTCTGTTTCAAGCGGTAGTGGAAATTTATCAGTAAAAAGATCGAGGGGCCCAAGGATTT 756
481 TTTCTGTTTCAAGCGGTAGTGGAAATTTATCAGTAAAAAGATCGAGGGGCCCAAGGATTT 540

QY 757 ACGCAACAACTTTAAGGCCCAATCTAGTGGAACTAGTTGTCCCAACATACGATGGAT 816
Db 541 ACGCAACAACTTTAAGGCCCAATCTAGTGGAACTAGTTGTCCCAACATACGATGGAT 600
QY 817 CCAAAATTTATGCCCGCTGATCCAGACTGCGATTTGCTTTATACATAGCAACATATTTGG 876
Db 601 CCAAAATTTATGCCCGCTGATCCAGACTGCGATTTGCTTTATACATAGCAACATATTTGG 660
QY 877 ATATCTAACTGTAACCCAGAGAAAGAGAGACTCACTTATGTGCACATAGGCTAGCC 936
Db 661 ATATCTAACTGTAACCCAGAGAAAGAGAGACTCACTTATGTGCACATAGGCTAGCC 720
QY 937 AACATGGAAGAGATGCCAGATCAGCTGGAGTCTCTTCTTCTCAAGAGAAATTT 996
Db 721 AACATGGAAGAGATGCCAGATCAGCTGGAGTCTCTTCTTCTCAAGAGAAATTT 780
QY 997 GATAGATTTCTGGCTATTTGGTGTGTCCTCAAGAGTCAAACTCCAGTGGTGTGTA 1056
Db 781 GATAGATTTCTGGCTATTTGGTGTGTCCTCAAGAGTCAAACTCCAGTGGTGTGTA 840
QY 1057 ATTTCTTAGAATTTCTATATGAAGAAATGATGAATCTGAGGTGGAATTTATTCATGTTACA 1116
Db 841 ATTTCTTAGAATTTCTATATGAAGAAATGATGAATCTGAGGTGGAATTTATTCATGTTACA 900
QY 1117 TCCCTTATGTTGAAACAAAGAGGAGGAGATTCATTCCTGTTATCTCTAAACAGGTACAGCA 1176
Db 901 TCCCTTATGTTGAAACAAAGAGGAGGAGATTCATTCCTGTTATCTCTAAACAGGTACAGCA 960
QY 1177 AATCTTAAAGTCACTTTTAAGATGTCAGAAATATGATGCTGGAAGAGGATCATATA 1236
Db 961 AATCTTAAAGTCACTTTTAAGATGTCAGAAATATGATGCTGGAAGAGGATCATATA 1020
QY 1237 GATGTCATAGATAAGGAACCTAATTCACCTTTTGGAGATTTCTATTTGAAGAGTGTGAATAT 1296
Db 1021 GATGTCATAGATAAGGAACCTAATTCACCTTTTGGAGATTTCTATTTGAAGAGTGTGAATAT 1080
QY 1297 ATTTGCCAGAGCTGGATGGATCTCTGAGGAAATATGCTTGGTCCATCTCTACTAGATCGC 1356
Db 1081 ATTTGCCAGAGCTGGATGGATCTCTGAGGAAATATGCTTGGTCCATCTCTACTAGATCGC 1140
QY 1357 TCCAGACTCGCTACAGATAGTGTGATCTCACTGAAATTTATTTATCCAGATGAGAAGAT 1416
Db 1141 TCCAGACTCGCTCGAGATAGTGTGATCTCACTGAAATTTATTTATCCAGATGAGAAGAT 1200
QY 1417 GATGTTATGAAAGGAGAGAGACTCATTTGAGTCACTGCTGATTTCTGTGACGCCACTAAT 1476
Db 1201 GATGTTATGAAAGGAGAGAGACTCATTTGAGTCACTGCTGATTTCTGTGAGGCCACTAAT 1260
QY 1477 ATCTATGAAGAAACAAACAGACATCTGGATTAATATCCATGACATCTTTTCTGTTTTCCC 1536
Db 1261 ATCTATGAAGAAACAAACAGACATCTGGATTAATATCCATGACATCTTTTCTGTTTTCCC 1320
QY 1537 CAAGTCAAGAGAGAAATTTGAGTTATTTTGGCTCTGATGCAATGCAAAACAGGTTTCCGT 1596
Db 1321 CAAGTCAAGAGAGAAATTTGAGTTATTTTGGCTCTGATGCAATGCAAAACAGGTTTCCGT 1380
QY 1597 CATTTATACAAATTTACATCTATTTTAAAGGAAAGCAATATAAACCAGTCCAGTGGTGG 1656
Db 1381 CATTTATACAAATTTACATCTATTTTAAAGGAAAGCAATATAAACCAGTCCAGTGGTGG 1440
QY 1657 CTGCTCTCTCAAGTGTCAAGTGTCTATCAAGAGGAGATAGCAATTTACCAGTGGT 1716
Db 1441 CTGCTCTCTCAAGTGTCAAGTGTCTATCAAGAGGAGATAGCAATTTACCAGTGGT 1500
QY 1717 GAATGGGAAGTTCTTGGCCGAGTGAATTAATATCCAGTGTGATGAAAGTCAAGAGGCTG 1776
Db 1501 GAATGGGAAGTTCTTGGCCGAGTGAATTAATATCCAGTGTGATGAAAGTCAAGAGGCTG 1560
QY 1777 GTATATTTTGAAGGACCAAAAGACTCCCTTTTAGAGCATCACTGTACGTAGTCAAGTTAC 1836
Db 1561 GTATATTTTGAAGGACCAAAAGACTCCCTTTTAGAGCATCACTGTACGTAGTCAAGTTAC 1620
QY 1837 GTAAATCTCGAGAGGTGACAAGGCTGACTGACCCGTGCTACTCAATTTCTTGTGCTGATC 1896

Db	1621	GTAAATCCTGGAGAGTGAACAGGCTGAACCGTGGCTACTCAATTTCTTGTGTGATC	1680
Qy	1897	AGTCAGCACTGTGAATCTTTTATAAGTAAGTATAGTAACCAAGAATCCACACTGTGTG	1956
Db	1681	AGTCAGCACTGTGAATCTTTTATAAGTAAGTATAGTAACCAAGAATCCACACTGTGTG	1740
Qy	1957	TCCTCTTTACAAGCTATCAAGTCCTGAAGATGACCCAACTTGGAAAACAAAGAAATTTTGG	2016
Db	1741	TCCTCTTTACAAGCTATCAAGTCCTGAAGATGACCCAACTTGGAAAACAAAGAAATTTTGG	1800
Qy	2017	GCACCAATTTTGGATTCAGCAGGTCCTCTTCTCTGACTATACCTCTCCAGAAAATTTTCTCT	2076
Db	1801	GCACCAATTTTGGATTCAGCAGGTCCTCTTCTCTGACTATACCTCTCCAGAAAATTTTCTCT	1860
Qy	2077	TTTGAAGTACTACTTGGATTTACATCTGATGGGATGCTCTACAAGSCTCATGATCTACAG	2136
Db	1861	TTTGAAGTACTACTTGGATTTACATCTGATGGGATGCTCTACAAGSCTCATGATCTACAG	1920
Qy	2137	CCTGGAAAATAATCTCTACTGTGCTGCTTCATATATGGTGGTCTCTCAGGTGCGATGGTG	2196
Db	1921	CCTGGAAAATAATCTCTACTGTGCTGCTTCATATATGGTGGTCTCTCAGGTGCGATGGTG	1980
Qy	2197	AATAATCGGTTTTAAAGAGAGTCAAGTATTTCCGCTTGGATACCCCTAGCTCTCTCAGGTTAT	2256
Db	1981	AATAATCGGTTTTAAAGAGAGTCAAGTATTTCCGCTTGGATACCCCTAGCTCTCTCAGGTTAT	2040
Qy	2257	GTGGTTGTAGTAGACACAACAGGGATCCTGTCAACGAGGGCTTAAATTTGAAGGGCC	2316
Db	2041	GTGGTTGTAGTAGACACAACAGGGGATCCTGTCAACGAGGGCTTAAATTTGAAGGGCC	2100
Qy	2317	TTTAAATATAAAATGGGTCAAATAGAAATGACGATCAGGTGGAGAGCTCCAAATATCTA	2376
Db	2101	TTTAAATATAAAATGGGTCAAATAGAAATGACGATCAGGTGGAGAGCTCCAAATATCTA	2160
Qy	2377	GCTTCTCGATGATGATTCATGTGCTTAGATCTGTGGGCATCCACGGCTGGTCTATGGA	2436
Db	2161	GCTTCTCGATGATGATTCATGTGCTTAGATCTGTGGGCATCCACGGCTGGTCTATGGA	2220
Qy	2437	GGATACCTCTCCGTATGGGCATTAATGACAGGTCAGATATCTTCAGGGTTGCTATTGCT	2496
Db	2221	GGATACCTCTCCGTATGGGCATTAATGACAGGTCAGATATCTTCAGGGTTGCTATTGCT	2280
Qy	2497	GGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAACTTATATGGGT	2556
Db	2281	GGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAACTTATATGGGT	2340
Qy	2557	CACCTTGACCAAGATGAACAGGGCTATTACTTGGATCTGTGGCCATGCAAGCAGAAAAG	2616
Db	2341	CACCTTGACCAAGATGAACAGGGCTATTACTTGGATCTGTGGCCATGCAAGCAGAAAAG	2400
Qy	2617	TTCCCTCTGAAACAAATCGTTTACTGCTCTTACATGTTTTCTCTGGATGAGAATGTCAT	2676
Db	2401	TTCCCTCTGAAACAAATCGTTTACTGCTCTTACATGTTTTCTCTGGATGAGAATGTCAT	2460
Qy	2677	TTTGACATACCAAGTATATTACTGAGTTTTTTTGTGAGGGCTGGAAGCCATATGATTTA	2736
Db	2461	TTTGACATACCAAGTATATTACTGAGTTTTTTTGTGAGGGCTGGAAGCCATATGATTTA	2520
Qy	2737	CAGATCTATCCTCAGGAGAGACACAGCATTAAGAGTTCTGAACTCGGAGAACATTATGAA	2796
Db	2521	CAGATCTATCCTCAGGAGAGACACAGCATTAAGAGTTCTGAACTCGGAGAACATTATGAA	2580
Qy	2797	CTGCATCTTTTGCATCTACCTTCAAGAAAACCTTGGATCAGTATTCCTCTCTCAAAAGTG	2856
Db	2581	CTGCATCTTTTGCATCTACCTTCAAGAAAACCTTGGATCAGTATTCCTCTCTCAAAAGTG	2640
Qy	2857	ATA 2859	
Db	2641	ATA 2643	

RESULT 8

ABX12255	
ID	ABX12255 standard; cDNA; 2649 BP.
XX	
AC	ABX12255;
XX	
DT	19-MAY-2003 (first entry)
XX	
DE	cDNA encoding human serine protease HIPHUM46.

Human; ss; gene; HIPHUM46; serine protease

serine protease activity modulation; diff
musculoskeletal disease; Hepatitis B vir

amyotrophic lateral sclerosis; paraneoplastic disease; paraproteinemia

irritable bowel syndrome; type
haemorrhoid; proctitis; rect.

colorectal tumour; cerebrovascular disease; multiple sclerosis; chronic

Homo sapiens.

Key	Location
1	1
2	2
3	3
4	4
5	5
6	6
7	7
8	8
9	9
10	10
11	11
12	12
13	13
14	14
15	15
16	16
17	17
18	18
19	19
20	20
21	21
22	22
23	23
24	24
25	25
26	26
27	27
28	28
29	29
30	30
31	31
32	32
33	33
34	34
35	35
36	36
37	37
38	38
39	39
40	40
41	41
42	42
43	43
44	44
45	45
46	46
47	47
48	48
49	49
50	50
51	51
52	52
53	53
54	54
55	55
56	56
57	57
58	58
59	59
60	60
61	61
62	62
63	63
64	64
65	65
66	66
67	67
68	68
69	69
70	70
71	71
72	72
73	73
74	74
75	75
76	76
77	77
78	78
79	79
80	80
81	81
82	82
83	83
84	84
85	85
86	86
87	87
88	88
89	89
90	90
91	91
92	92
93	93
94	94
95	95
96	96
97	97
98	98
99	99
100	100

33

```

/product= "H1PHUM46"
/note= "Serine protease"

```

GB2374869-A.

30-OCT-2002, 22-JAN-2002, 2002CH-00001404

23-TAN-2001: 2001CB-00001750

(GLAX) GLAYO GROUP LTD

Matches 2646;		Conservative	0;	Mismatches	3;	Indels	0;	Gaps	0;
QY	214	ATGGCAGCAGCAAT	TGGAAACAGAACACGCTGGGTGTTGAGATATTTGAAATCGCGACTGT	273					
Db	1	ATGGCAGCAGCAAT	TGGAAACAGAACACGCTGGGTGTTGAGATATTTGAAATCGCGACTGT	60					
QY	274	GAGGAGATATTTGAAT	CACAGGATCGGCCCTAAATTTGGAGCCTTTTATGTTGAGCGGTAT	333					
Db	61	GAGGAGATATTTGAAT	CACAGGATCGGCCCTAAATTTGGAGCCTTTTATGTTGAGCGGTAT	120					
QY	334	TCCTGGAGTCAGCTTAAAGCT	TGCGCATGCGGATACAGAAATATACATGCTTACATGATG	393					
Db	121	TCCTGGAGTCAGCTTAAAGCT	TGCGCATGCGGATACAGAAATATACATGCTTACATGATG	180					
QY	394	GCTAAGGCACACATGATTTCA	TGTTGTGAAAGAGAAATGATCCAGATGGACCTCATTC	453					
Db	181	GCTAAGGCACACATGATTTCA	TGTTGTGAAAGAGAAATGATCCAGATGGACCTCATTC	240					
QY	454	GACGAATCTATTACCTTGCC	ATGCTCTGAGAGAAATACACTGTTTATTTCT	513					
Db	241	GACGAATCTATTACCTTGCC	ATGCTCTGAGAGAAATACACTGTTTATTTCT	300					
QY	514	GAAATTTCCAAAATATCAAT	ATAGACAGCAGTCTTAATGCTCTCTTGGAAAGCCTCTTTTG	573					
Db	301	GAAATTTCCAAAATATCAAT	ATAGACAGCAGTCTTAATGCTCTCTTGGAAAGCCTCTTTTG	360					
QY	574	GATCTTTTTCAGGCAACACT	GGACTATGGAATGTATTTCTCGAGAGAAAGAACTATTAA	633					
Db	361	GATCTTTTTCAGGCAACACT	GGACTATGGAATGTATTTCTCGAGAGAAAGAACTATTAA	420					
QY	634	GAAAGAAACGATTTGGAACA	GTCGGAAATTTCTTACGATTTATCAACAAAGGAAGTGA	693					
Db	421	GAAAGAAACGATTTGGAACA	GTCGGAAATTTCTTACGATTTATCAACAAAGGAAGTGA	480					
QY	694	ACATTTCTGTTTCAAGCGGT	AGTGGAAATTTATCAAGTAAAGATGGAGGGCCACAAGGA	753					
Db	481	ACATTTCTGTTTCAAGCGGT	AGTGGAAATTTATCAAGTAAAGATGGAGGGCCACAAGGA	540					
QY	754	TTTAGGCAACACCTTTTAA	GCCCACTAGTGGAAACTAGTTGTCCTCCCAATACGAGATG	813					
Db	541	TTTAGGCAACACCTTTTAA	GCCCACTAGTGGAAACTAGTTGTCCTCCCAATACGAGATG	600					
QY	814	GATCCAAATTTATGCCCGT	GTATCCAGACTGATGCTTTTATACATAGCAACGATATT	873					
Db	601	GATCCAAATTTATGCCCGT	GTATCCAGACTGATGCTTTTATACATAGCAACGATATT	660					
QY	874	TGATATCTAATCATCGTAA	CCAGAGAAAGAGAGACTCACTTATGTCACAATGAGCTA	933					
Db	661	TGATATCTAATCATCGTAA	CCAGAGAAAGAGAGACTCACTTATGTCACAATGAGCTA	720					
QY	934	GCACAATCGGAAGAGATGC	CAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAAAGAGAA	993					
Db	721	GCACAATCGGAAGAGATGC	CAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAAAGAGAA	780					
QY	994	TTTGATAGATATCTGGCTA	TTTGGTGTCCAAAAGCTGAAACAACTCCCAAGTGGTGGT	1053					
Db	781	TTTGATAGATATCTGGCTA	TTTGGTGTCCAAAAGCTGAAACAACTCCCAAGTGGTGGT	840					
QY	1054	AAATTTCTTAGAATCTTAT	ATGAAAGAAATGATGAATCTGAGGTGGAAATTTATCATGTT	1113					
Db	841	AAATTTCTTAGAATCTTAT	ATGAAAGAAATGATGAATCTGAGGTGGAAATTTATCATGTT	900					
QY	1114	ACATCCCTATGTTGGAACA	CAAGAGGGCAGATTCATTCGGTTATCTTAAACAGGTACA	1173					
Db	901	ACATCCCTATGTTGGAACA	CAAGAGGGCAGATTCATTCGGTTATCTTAAACAGGTACA	960					
QY	1174	GCAAATCTTAAAGTCAC	TTTTAAGATGTCAGAAATTAATGATTTGATGCTGGAAGGAAGATC	1233					
Db	961	GCAAATCTTAAAGTCAC	TTTTAAGATGTCAGAAATTAATGATTTGATGCTGGAAGGAAGATC	1020					
QY	1234	ATAGATGTCATAGATAAG	GAACCTAAATCAACCTTTTGGAGATTTCTATTTGAAGAGGTTGAA	1293					
Db	1021	ATAGATGTCATAGATAAG	GAACCTAAATCAACCTTTTGGAGATTTCTATTTGAAGAGGTTGAA	1080					

QY	1294	TATATTGCCAGAGCTGGAT	GGACTCCTGTGAGGAAAAATATGCTTGGTCCATCCTACTAGAT	1353					
Db	1081	TATATTGCCAGAGCTGGAT	GGACTCCTGTGAGGAAAAATATGCTTGGTCCATCCTACTAGAT	1140					
QY	1354	CGCTCCAGAGCTCGCCTTA	CAGATAGTGTGATCTCACTGGAATTAATTTATCCAGTAGAA	1413					
Db	1141	CGCTCCAGAGCTCGCCTTA	CAGATAGTGTGATCTCACTGGAATTAATTTATCCAGTAGAA	1200					
QY	1414	GATGATGTTATGGAAGGCA	GCAGAGACTCAATGAGTCAGTCGCTGATCTGTGACGCCACTA	1473					
Db	1201	GATGATGTTATGGAAGGCA	GCAGAGACTCAATGAGTCAGTCGCTGATCTGTGACGCCACTA	1260					
QY	1474	ATTATCTATGAAAGAAACA	CAGACATCTGGATAAATATCCATGACATCTTTTCATGTTTTT	1533					
Db	1261	ATTATCTATGAAAGAAACA	CAGACATCTGGATAAATATCCATGACATCTTTTCATGTTTTT	1320					
QY	1534	CCCCAAAGTCACGAAGAG	GAATTTGAGTTTATTTTTTGGCTCTGAAATGCAAAAACAGGTTTC	1593					
Db	1321	CCCCAAAGTCACGAAGAG	GAATTTGAGTTTATTTTTTGGCTCTGAAATGCAAAAACAGGTTTC	1380					
QY	1594	CGTCATTTTATCAAAAAT	TACATCTATTTTAAAGGAAGCAAAATATAAACGATCCAGTGGT	1653					
Db	1381	CGTCATTTTATCAAAAAT	TACATCTATTTTAAAGGAAGCAAAATATAAACGATCCAGTGGT	1440					
QY	1654	GGGCTGCTGCTCCAAAGT	GTCTTATCAAAAGAGAGATAGCAATTAACAGT	1713					
Db	1441	GGGCTGCTGCTCCAAAGT	GTCTTATCAAAAGAGAGATAGCAATTAACAGT	1500					
QY	1714	GGTGAAATGGGAAGTTCT	TGGCGCGCATGGATCTAATATCCAAGTTGATGAAGTCAGAAGG	1773					
Db	1501	GGTGAAATGGGAAGTTCT	TGGCGCGCATGGATCTAATATCCAAGTTGATGAAGTCAGAAGG	1560					
QY	1774	CTGGTATATTTTGAAGGC	ACCAAGACTCCCTTTAGAGCATCACTGTACGTAGTCAGT	1833					
Db	1561	CTGGTATATTTTGAAGGC	ACCAAGACTCCCTTTAGAGCATCACTGTACGTAGTCAGT	1620					
QY	1834	TACGTAATCTCGAGAGGT	GACAAGCTGACCTGACCTGGCTACTCACATTTCTTGCTGC	1893					
Db	1621	TACGTAATCTCGAGAGGT	GACAAGCTGACCTGACCTGGCTACTCACATTTCTTGCTGC	1680					
QY	1894	ATCAGTCAGCACTGTGACT	CTTTTATAAAGTATAGTAACCAAGAAATCCACACTGT	1953					
Db	1681	ATCAGTCAGCACTGTGACT	CTTTTATAAAGTATAGTAACCAAGAAATCCACACTGT	1740					
QY	1954	GTGTCCCTTTTAAAGCT	ATCAAGTCTCGAAGTACCCAACTTGCAAAACAAAGGAAATTT	2013					
Db	1741	GTGTCCCTTTTAAAGCT	ATCAAGTCTCGAAGTACCCAACTTGCAAAACAAAGGAAATTT	1800					
QY	2014	TGGGCGCACCATTTTGGAT	TCAGCAGGTCTCTTCTCTGACTATATCTCTCCAGAAAATTTTC	2073					
Db	1801	TGGGCGCACCATTTTGGAT	TCAGCAGGTCTCTTCTCTGACTATATCTCTCCAGAAAATTTTC	1860					
QY	2074	TCCTTTGAAAGTACTACT	CGGATTTACATGTATGGGATGCTCTACAGCCCTCATGATCTA	2133					
Db	1861	TCCTTTGAAAGTACTACT	CGGATTTACATGTATGGGATGCTCTACAGCCCTCATGATCTA	1920					
QY	2134	CAGCTCGAAAGAAATAT	CTCTGCTGCTTATATATGGTGGTCTCTCAGGTGAGTTG	2193					
Db	1921	CAGCTCGAAAGAAATAT	CTCTGCTGCTTATATATGGTGGTCTCTCAGGTGAGTTG	1980					
QY	2194	GTGAATATCGGTTTAAAG	GAGTCAAGTATTTCCGCTTGAATACCTTAGCCCTCTCTAGGT	2253					
Db	1981	GTGAATATCGGTTTAAAG	GAGTCAAGTATTTCCGCTTGAATACCTTAGCCCTCTCTAGGT	2040					
QY	2254	TATGTGGTTGTAGTAGA	CAACAGGGATCTGTCAACCGAGGGCTTAAATTTGAAGGC	2313					
Db	2041	TATGTGGTTGTAGTAGA	CAACAGGGATCTGTCAACCGAGGGCTTAAATTTGAAGGC	2100					
QY	2314	GCCTTTAAATATAAAT	TGGGTCAAAATAGAAATTTGACATCAGGTGGAGGACTCCAAATAT	2373					
Db	2101	GCCTTTAAATATAAAT	TGGGTCAAAATAGAAATTTGACATCAGGTGGAGGACTCCAAATAT	2160					

QY 2374 CTAGCTTCTCGATATGATTTTCATTTAGCTTAGATCGTGTGGGATCCAGCGGTGGTCTCTAT 2433
Db |||||
QY 2434 GGAGATACCTCTCCCTGATGGCATTATGTCAGAGTCAAGATATCTTCAGGGTTCCTATT 2493
Db |||||
QY 2494 GCTGGGGCCCCAGTCACTCTGTGTGATCTTCTATGATACAGGATACACGGAACGTTATATG 2553
Db |||||
QY 2554 GGTCACTCCGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCCATCAACAGAGAA 2613
Db |||||
QY 2614 AAGTTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGTTTCTCGATGAGATGTC 2673
Db |||||
QY 2674 CATTTTGCACATACAGTATATTACTGAGTGTCTTCTAGTGGGCTGGAAAGCCATATGAT 2733
Db |||||
QY 2734 TTACAGATCTATCTCTCAGGAGACACAGCATGAAGTTCCTGATCGGAGACATATTAT 2793
Db |||||
QY 2794 GNACTGCATCTTTTGGCACTACCTTCAAGAAACCTTGGATCAGTATTGCTGCTCTAAAA 2853
Db |||||
QY 2854 GTGATATATA 2862
Db |||||
QY 2862 GTGATATA 2649
Db |||||
RESULT 9
ID ABK83322 standard; cDNA; 2671 BP.
AC ABK83322;
DT 12-AUG-2002 (first entry)
DE cDNA encoding human DPP4V related serine protease DPP-1.
KW Human; serine protease; dipeptidyl peptidase IV-related protein; DPP;
KW DPP4V; infection; human immunodeficiency virus; HIV-1; HIV-2; pain;
KW diabetes; infertility; obesity; anorexia; Parkinson's disease; stroke;
KW heart failure; hypertension; urinary retention; osteoporosis; cancer;
KW ulcer; allergy; cancer; psychotic disorder; neurological disorder;
KW dyskinesia; reproductive disorder; inflammatory disorder;
KW metabolic disorder; gene; 86.
OS Homo sapiens.
XX
XX WO200231134-A2.
PN
PD 18-APR-2002.
XX
PF 12-OCT-2001; 2001WO-US031874.
XX
PR 12-OCT-2000; 2000US-0240117P.
XX
XX (FERR) FERRING BV.
PA
XX Qi S, Akinsanya KO, Riviere PJ, Junien J;
PI
XX WPI; 2002-444178/18.
DR P-PSDB; ABG61591.
XX
XX New dipeptidyl peptidase IV-related proteins and nucleic acids encoding

PT the proteins, useful for treating e.g. fungal, bacterial, protozoan and
Pt viral infections, cancers, allergies, neurological disorders, or pain.
XX
PS Claim 1; Page 53-54; 113pp; English.
XX The present invention relates to the isolation of novel human serine
CC proteases referred to as dipeptidyl peptidase IV (DPP4V)-related proteins
CC (DPPP). The dipeptidyl peptidase IV-related proteins (DPPP) and nucleic
CC acids encoding them are useful for treating infections such as fungal,
CC bacterial, protozoan and viral infections, particularly infections caused
CC by human immunodeficiency virus (HIV-1 or HIV-2), pain, diabetes,
CC precocious puberty, infertility, obesity, anorexia, bulimia, Parkinson's
CC disease, acute heart failure, hypotension, hypertension, urinary
CC retention, osteoporosis, angina pectoris, psychotic and neurological
CC allergies, cancers, migraine, vomiting, stroke, ulcers, asthma,
CC disorders (e.g. anxiety, dementia, or schizophrenia), and dyskinesias.
CC These may also be used in discovering therapeutic agents for the
CC treatment of reproductive, inflammatory and metabolic disorders. ABK83322
CC -ABK83343 encode human DPPP proteins
XX
SQ Sequence 2671 BP; 805 A; 524 C; 594 G; 748 T; 0 U; 0 Other;
Query Match 80.0%; Score 2495; DB 6; Length 2671;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2645; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 213 CATGGCAGCAGCAATGGAAACAGAACAGCTGGGTGTGAGATATTTGAACTCGGACTG 272
Db |||||
QY 7 CATGGCAGCAGCAATGGAAACAGAACAGCTGGGTGTGAGATATTTGAACTCGGACTG 66
Db |||||
QY 273 TGAGGAAATATTGAATCAAGATCGGCTAAATTTGGAGCCCTTTTATGTTGAGCGGTA 332
Db |||||
QY 67 TGAGGAAATATTGAATCAAGATCGGCTAAATTTGGAGCCCTTTTATGTTGAGCGGTA 126
Db |||||
QY 333 TTCCTGAGTCAAGTTTAAAGCTGCTTCCGATACCGAAATATCATGCTACATGAT 392
Db |||||
QY 127 TTCCTGAGTCAAGTTTAAAGCTGCTTCCGATACCGAAATATCATGCTACATGAT 186
Db |||||
QY 393 GGCTAAGGCCACCATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGGACCTCATTC 452
Db |||||
QY 187 GGCTAAGGCCACCATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGGACCTCATTC 246
Db |||||
QY 453 AGACAGAAATCTATTACCTTGGCATGCTGTGAGAACACAGAGAAATACACCTGTTTATTC 512
Db |||||
QY 247 AGACAGAAATCTATTACCTTGGCATGCTGTGAGAACACAGAGAAATACACCTGTTTATTC 306
Db |||||
QY 513 TGAATTTCCCAAACTATCAATAGACGACGCTCTTAATGCTCTCTTGGAGCCCTCTTTT 572
Db |||||
QY 307 TGAATTTCCCAAACTATCAATAGACGACGCTCTTAATGCTCTCTTGGAGCCCTCTTTT 366
QY 573 GGATCTTTTTCAGGCAACACTGGACTATGATGATGTTCTCGAGAGAGAACTATTAAAG 632
Db |||||
QY 367 GGATCTTTTTCAGGCAACACTGGACTATGATGATGTTCTCGAGAGAGAACTATTAAAG 426
QY 633 AGAAGAAAACCGATTTGGAACAGTGGGATTCCTTACGATTTATCACCAGGAGTGG 692
Db |||||
QY 427 AGAAGAAAACCGATTTGGAACAGTGGGATTCCTTACGATTTATCACCAGGAGTGG 486
QY 693 AACATTTCTGTTTCAAGCCGGTAGTGAATTTATCAGTAAAGATGAGGGGCCACAAAGG 752
Db |||||
QY 487 AACATTTCTGTTTCAAGCCGGTAGTGAATTTATCAGTAAAGATGAGGGGCCACAAAGG 546
QY 753 ATTTACGCAACAACTTTTAAAGGCCCAATCTAGTGAATCTAGTGTGTTCCCAACATACCGAT 812
Db |||||
QY 547 ATTTACGCAACAACTTTTAAAGGCCCAATCTAGTGAATCTAGTGTGTTCCCAACATACCGAT 606
QY 813 GGATCCAAAATTTATGCCCGCTGATCCAGACTGGATTCCTTTATACATACCAACGATAT 872
Db |||||
QY 607 GGATCCAAAATTTATGCCCGCTGATCCAGACTGGATTCCTTTATACATACCAACGATAT 666
QY 873 TTGATATCTAACATCGTAACAGAGAGAAAGAGAGCTCACTTATGTCACAAATGAGCT 932
Db |||||
QY 667 TTGATATCTAAACATCGTAACAGAGAGAAAGAGAGCTCACTTATGTCACAAATGAGCT 726
Db |||||

```
QY 933 AGCAACATGGAGAGATGCCAGATCAGCTGGAGTGGTCTTCTCCCAAGA 992
DB 727 AGCAACATGGAGAGATGCCAGATCAGCTGGAGTGGTCTTCTCCCAAGA 786
QY 993 APTTGTAGATATTCCTGGCTATTTGGTGTCCAAAGCTGAAACAACTCCCAAGTGGTG 1052
DB 787 ATTGTAGATATTCCTGGCTATTTGGTGTCCAAAGCTGAAACAACTCCCAAGTGGTG 846
QY 1053 TAAAAATCTTAGAATCTATATCAAGAAAAATGATGAATCTGAGTGGAAAAATTAATCATGT 1112
DB 847 TAAAAATCTTAGAATCTATATGAAGAAAAATGATGAATCTGAGTGGAAAAATTAATCATGT 906
QY 1113 TACATCCCTATGTTGGAAAAAGAGGAGGAGATTCATTCGGTTATCCTAAAAAGGTAC 1172
DB 907 TACATCCCTATGTTGGAAAAAGAGGAGGAGATTCATTCGGTTATCCTAAAAAGGTAC 966
QY 1173 AGCAATCTTAAAGTCACTTTTAAAGATGTCAGAAATATGATGATCTGATCTGAAGAGGAT 1232
DB 967 AGCAATCTTAAAGTCACTTTTAAAGATGTCAGAAATATGATGATCTGATCTGAAGAGGAT 1026
QY 1233 CATAGATGTCATAGATAAGGACTAATTCACCTTTTTCAGATCTCTATTTGAAGAGTTGA 1292
DB 1027 CATAGATGTCATAGATAAGGACTAATTCACCTTTTTCAGATCTCTATTTGAAGAGTTGA 1086
QY 1293 ATATATTTGCGAGCTGGATGACCTCTGAGGAAAAATATCTTGGTCCATCTTACTAGA 1352
DB 1087 ATATATTTGCGAGCTGGATGACCTCTGAGGAAAAATATCTTGGTCCATCTTACTAGA 1146
QY 1353 TCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAAATTAATTTATCCAGPAGA 1412
DB 1147 TCGCTCCAGACTCGCTCGAGATAGTGTGATCTCACCTGAAATTAATTTATCCAGPAGA 1206
QY 1413 AGATAGTGTATGGAAGGAGAGACTCAATGAGTCACTGAGTCACTGAGTCACTGAGTCACT 1472
DB 1207 AGATAGTGTATGGAAGGAGAGACTCAATGAGTCACTGAGTCACTGAGTCACTGAGTCACT 1266
QY 1473 AATTATCTATGAGAAAAACAAGACATCTGGATTAATATCCATGACATCTTTCATGTTTT 1532
DB 1267 AATTATCTATGAGAAAAACAAGACATCTGGATTAATATCCATGACATCTTTCATGTTTT 1326
QY 1533 TCCCAAGTCAAGAGAGAAATTCAGTTTATTTTCTGCTGAAATCAAAACAGGTTT 1592
DB 1327 TCCCAAGTCAAGAGAGAAATTCAGTTTATTTTCTGCTGAAATCAAAACAGGTTT 1386
QY 1593 CCGTCAATTTATACAAATATACATCTATTTTAAAGGAAAGCAATATAACATCCAGTGG 1652
DB 1387 CCGTCAATTTATACAAATATACATCTATTTTAAAGGAAAGCAATATAACATCCAGTGG 1446
QY 1653 TGGCTGCTGCTCCAGTGAATTTCAAGTGTCTATCAAGAGGAGATAGCAATTAACAG 1712
DB 1447 TGGCTGCTGCTCCAGTGAATTTCAAGTGTCTATCAAGAGGAGATAGCAATTAACAG 1506
QY 1713 TGGTGAATGGGAAGTTCTTGGCGGAGATGATCTAATATCCAAAGTTGATGAGTCAAG 1772
DB 1507 TGGTGAATGGGAAGTTCTTGGCGGAGATGATCTAATATCCAAAGTTGATGAGTCAAG 1566
QY 1773 GCTGGTATATTTTGAAGGACCAAGACTCCCTTTTGGAGCATCACTGACGTAGTACAG 1832
DB 1567 GCTGGTATATTTTGAAGGACCAAGACTCCCTTTTGGAGCATCACTGACGTAGTACAG 1626
QY 1833 TTACGTAATCTCGAGAGGTGACAGGCTGACCTGCTGCTACTCACATTTCTTGCTG 1892
DB 1627 TTACGTAATCTCGAGAGGTGACAGGCTGACCTGCTGCTACTCACATTTCTTGCTG 1686
QY 1893 CATCAGTCAGACTGTGACTTTCTTTATAGTAAGTATAGTAACCAAGAAATCCACACTG 1952
DB 1687 CATCAGTCAGACTGTGACTTTCTTTATAGTAAGTATAGTAACCAAGAAATCCACACTG 1746
QY 1953 TGTGTCCCTTTTACAGCTTATCAAGTCTGGAAGATGACCAACTTGTGAAAAACAAGGAATT 2012
DB 1747 TGTGTCCCTTTTACAGCTTATCAAGTCTGGAAGATGACCAACTTGTGAAAAACAAGGAATT 1806
```

```
QY 2013 TTGGGCCACCAATTTTGGATTTCAGAGGTCTCTTCTCTGACTATATCTCTCCAGAAATTTT 2072
DB 1807 TTGGGCCACCAATTTTGGATTTCAGAGGTCTCTTCTCTGACTATATCTCTCCAGAAATTTT 1866
QY 2073 CTCTTTTGAAGATCTACTGGAATTTACATTTATATGGGATGCTCTACAGGCTCATGATCT 2132
DB 1867 CTCTTTTGAAGATCTACTGGAATTTACATTTATATGGGATGCTCTACAGGCTCATGATCT 1926
QY 2133 ACAGCCTGGAAGAAATATCTACTGCTCTCTATATATGTTGCTCTCTCAGGTGCAAGTT 2192
DB 1927 ACAGCCTGGAAGAAATATCTACTGCTCTCTATATATGTTGCTCTCTCAGGTGCAAGTT 1986
QY 2193 GGTGAATAATTCGGTTTAAAGGAGTCAAGTATTTTCCGCTTGAATACCCCTAGCCTCTTAGG 2252
DB 1987 GGTGAATAATTCGGTTTAAAGGAGTCAAGTATTTTCCGCTTGAATACCCCTAGCCTCTTAGG 2046
QY 2253 TTATGTGGTTGTAGTGATAGACAAACAGGGGATCTCTGTACCGAGGGCTTAAATTTGAAGG 2312
DB 2047 TTATGTGGTTGTAGTGATAGACAAACAGGGGATCTCTGTACCGAGGGCTTAAATTTGAAGG 2106
QY 2313 CGCCTTTAAATATAAATGGGTCAAAATAGAAATTTGACGATCAGGTGGAAGACTCCAAATA 2372
DB 2107 CGCCTTTAAATATAAATGGGTCAAAATAGAAATTTGACGATCAGGTGGAAGACTCCAAATA 2166
QY 2373 TCTAGCTTCTCGATATGATTTCAITGACTTAGATCGTGTGGGCATCCACGGCTGGTCTTA 2432
DB 2167 TCTAGCTTCTCGATATGATTTCAITGACTTAGATCGTGTGGGCATCCACGGCTGGTCTTA 2226
QY 2433 TGGAGGATACCTCTCCCTGATGSCATTAATGCAGAGTTCAGATATCTTCAGGGTTGCTAT 2492
DB 2227 TGGAGGATACCTCTCCCTGATGSCATTAATGCAGAGTTCAGATATCTTCAGGGTTGCTAT 2286
QY 2493 TGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAACGTTATAT 2552
DB 2287 TGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAACGTTATAT 2346
QY 2553 GGGTCACCCCTGACAGAGATGAACAGGGCTTATTAATTAGGATCTGTGGCCATGCAAGCAGA 2612
DB 2347 GGGTCACCCCTGACAGAGATGAACAGGGCTTATTAATTAGGATCTGTGGCCATGCAAGCAGA 2406
QY 2613 AAAGTTCCCTCTGAAACCAAAATCGTTTACTGCTCTTACATGTTTCTTGGATGAGATGT 2672
DB 2407 AAAGTTCCCTCTGAAACCAAAATCGTTTACTGCTCTTACATGTTTCTTGGATGAGATGT 2466
QY 2673 CCATTTTGCACATACCAAGTATATTAATGATGTTTTTGTAGGGCTGAAAGCCATATGA 2732
DB 2467 CCATTTTGCACATACCAAGTATATTAATGATGTTTTTGTAGGGCTGAAAGCCATATGA 2526
QY 2733 TTTACAGATCTATCTCTCAGGAGACACAGCATAAAGATTCCTGAATCGGAGAACATTA 2792
DB 2527 TTTACAGATCTATCTCTCAGGAGACACAGCATAAAGATTCCTGAATCGGAGAACATTA 2586
QY 2793 TGAACCTGCATCTTTTTCGACTACCTTCAAGAAAAACCTTGGATCAAGTATGCTGCTTAAA 2852
DB 2587 TGAACCTGCATCTTTTTCGACTACCTTCAAGAAAAACCTTGGATCAAGTATGCTGCTTAAA 2646
QY 2853 AGTGATAT 2860
DB 2647 AGTGATAT 2654
```

RESULT 10
ADL13374

ID ADL13374 standard; cDNA; 2797 BP.

XX ADL13374;

XX 06-MAY-2004 (first entry)

DE Human steroid-induced C3A liver cell cDNA #1103.

XX ss; gene; Hepatotropic; Gene therapy; Wilson disease; liver disorder;
KW steroid therapy; cirrhosis; hepatitis; human; C3A liver cell.

XX OS Homo sapiens.
XX PN US6673549-B1.
XX PD 06-JAN-2004.
XX PF 12-OCT-2001; 2001US-00976594.
XX PR 12-OCT-2000; 2000US-0240409P.
XX PA (INCY-) INCYTE CORP.
XX PI Furness LM, Buchbinder JL;
XX DR WPI; 2004-068610/07.
XX PT Combination useful for preparing a composition for treating liver
XX PT disorders associated with steroid therapy, e.g., cirrhosis or hepatitis,
XX PT comprises cDNAs that are differentially expressed in response to steroid
XX PT treatment.
XX PS Claim 1; SEQ ID NO 1103; 141pp; English.

CC The invention relates to a combination comprising cDNAs that are
CC differentially expressed in response to steroid treatment. Also included
CC are the following: a high throughput method for using a cDNA to detect
CC differential expression of nucleic acids in a sample; and a high
CC throughput method of screening molecules or compounds to identify a
CC ligand that specifically binds a cDNA. The sample is from a subject with
CC Wilson disease and comparison of a standard defines a stage of that
CC disease. The high throughput method of screening molecules or compounds
CC to identify a ligand that specifically binds a cDNA comprises: combining
CC the combination with molecules or compounds under conditions to allow
CC specific binding; and detecting specific binding between each cDNA and at
CC least one molecule or compound. The molecules or compounds are regulatory
CC proteins. The combination is useful for preparing a composition for
CC treating liver disorders associated with steroid therapy, e.g., cirrhosis
CC or hepatitis. The present sequence represents a human cDNA which is
CC differentially expressed in steroid-induced C3A liver cells. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format directly from USPTO
CC at seqdata.uspto.gov/sequence.html.

XX SQ Sequence 2797 BP; 866 A; 551 C; 595 G; 785 T; 0 U; 0 Other;
Query Match 74.2%; Score 2315; DB 12; Length 2797;
Best Local Similarity 99.8%; Pred No. 0;
Matches 2785; Conservative 0; Mismatches 1; Indels 5; Gaps 2;

Qy 335 CCTGGAGTCAGCTTAAAGCTGCTTGGCGATACCAGAAATATCATGGCTACATGATGG 394
Db 1 CCTGGAGTCAGCTTAAAGCTGCTTGGCGATACCAGAAATATCATGGCTACATGATGG 60
Qy 395 CTAAGGCACCATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGGACCTCATTCAG 454
Db 61 CTAAGGCACCATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGGACCTCATTCAG 120
Qy 455 ACAGATCTATTAACCTTGCCATGTCGTGTGAGAAACAGAGAAATACACTGTTTTATCTG 514
Db 121 ACAGATCTATTAACCTTGCCATGTCGTGTGAGAAACAGAGAAATACACTGTTTTATCTG 180
Qy 515 AAATCCCAAACTATCAATAGACGACGATCTTAATGCTCTCTTGGAGCCCTCTTTTGG 574
Db 181 AAATCCCAAACTATCAATAGACGACGATCTTAATGCTCTCTTGGAGCCCTCTTTTGG 240
Qy 575 ATCTTTTTCAGGCAACACTGGACTATGGAATGATTTCTCGAAGAGGAAGCACTATTAAAG 634
Db 241 ATCTTTTTCAGGCAACACTGGACTATGGAATGATTTCTCGAAGAGGAAGCACTATTAAAG 300
Qy 635 AAAGAAAACGCATTGGAAACAGTCGGAATTGCTTCTTACGATTATCACCAGGAGTGGAA 694
Db 301 AAAGAAAACGCATTGGAAACAGTCGGAATTGCTTCTTACGATTATCACCAGGAGTGGAA 360

Qy 695 CATTTCTGTTTCAAGCCGCTAGTGGAAATTTATCACTAAAGATGAGGGGCCACAAGGAT 754
Db 361 CATTTCTGTTTCAAGCCGCTAGTGGAAATTTATCAGTAAAGATGAGGGGCCACAAGGAT 420
Qy 755 TTACGCAACACCTTTAAGGCCCAATCTAGTGGAAACTAGTTGTCCCAACATACGATGG 814
Db 421 TTACGCAACACCTTTAAGGCCCAATCTAGTGGAAACTAGTTGTCCCAACATACGATGG 480
Qy 815 ATCCAAAATTAATGCCCCGCTGATCCAGATCGGATTTGCTTTTATACATAGCAACGATATTT 874
Db 481 ATCCAAAATTAATGCCCCGCTGATCCAGATCGGATTTGCTTTTATACATAGCAACGATATTT 540
Qy 875 GGATATCTAAACATCGTAACCCAGAGAAAGAGAGACTCACTTATGTGCACAAATGAGCTAG 934
Db 541 GGATATCTAAACATCGTAACCCAGAGAAAGAGAGACTCACTTATGTGCACAAATGAGCTAG 600
Qy 935 CCAACATGGAAGAGATGCCAGATCAGCTGAGTGCCTACCTTTGTTTCTCCAAAGAGAAT 994
Db 601 CCAACATGGAAGAGATGCCAGATCAGCTGAGTGCCTACCTTTGTTTCTCCAAAGAGAAT 660
Qy 995 TTGATAGATATCTGGCTATTGGTGTGTCCTCAAAAAGCTGAAAACAACCTCCCAAGTGGTGA 1054
Db 661 TTGATAGATATCTGGCTATTGGTGTGTCCTCAAAAAGCTGAAAACAACCTCCCAAGTGGTGA 720
Qy 1055 AAATTTCTAGAAATCTATATGAAGAAAATGATGAATCTGAGTGGAATTTATTCATCTTA 1114
Db 721 AAATTTCTAGAAATCTATATGAAGAAAATGATGAATCTGAGTGGAATTTATTCATCTTA 780
Qy 1115 CATCCCCCTATGTTGGAACAAGG-AGGCGACAGATTCATTCCTGTTATCTCTAAACAGGTACA 1173
Db 781 CATCCCCCTATGTTGGAACAAGGCGAGGCGAGNATTCCTCGTTATCTCTAAACAGGTACA 840
Qy 1174 GCAAACTCTAAAGTCACTTTTAAAGATGTCAGAAATGATGATGATGCTGGAAGGAGATC 1233
Db 841 GCAAACTCTAAAGTCACTTTTAAAGATGTCAGAAATGATGATGATGCTGGAAGGAGATC 900
Qy 1234 ATAGATGTCATAGATGAAGAACTAATTCACCTTTGAGATCTTATTTGAAGGATTTGAA 1293
Db 901 ATAGATGTCATAGATGAAGAACTAATTCACCTTTGAGATCTTATTTGAAGGATTTGAA 960
Qy 1294 TATATTGCCAGAGCTGGATGGACTCTCTGAGGGAATAA- - - TGCTTGGTCCATCTCTACT 1349
Db 961 TATATTGCCAGAGCTGGATGGACTCTCTGAGGGAATAAAGTAGTGTCTTGGTCCATCTCTACT 1020
Qy 1350 AGATCGCTCCAGACTCGCCTACAGATAGTGTGATCTCACCTGAAATTTATTTATCCAGT 1409
Db 1021 AGATCGCTCCAGACTCGCCTACAGATAGTGTGATCTCACCTGAAATTTATTTATCCAGT 1080
Qy 1410 AGAAGATGATGTTATGGAAGGCGAGACTCAATTGAGTCACTGCTGATCTGTGACGCC 1469
Db 1081 AGAAGATGATGTTATGGAAGGCGAGACTCAATTGAGTCACTGCTGATCTGTGACGCC 1140
Qy 1470 ACTAATTTATCTATGAAGAAACAAACAGACATCTGGATAAATATCCATGACATCTTTCACTGT 1529
Db 1141 ACTAATTTATCTATGAAGAAACAAACAGACATCTGGATAAATATCCATGACATCTTTCACTGT 1200
Qy 1530 TTTTCCCAAAAGTCACGAAGAGGAAATTTAGTGTATTTTTCCTCTGAAATGCAAAACAGG 1589
Db 1201 TTTTCCCAAAAGTCACGAAGAGGAAATTTAGTGTATTTTTCCTCTGAAATGCAAAACAGG 1260
Qy 1590 TTTTCCGTCATTTATACAAATTTACATCTATTTTAAAGGAAGCAATATATACGATCCAG 1649
Db 1261 TTTTCCGTCATTTATACAAATTTACATCTATTTTAAAGGAAGCAATATATACGATCCAG 1320
Qy 1650 TGGTGGCTGCTGCTCCAGTGATTTTCAGTGCTCTATCAAGAGAGAGTAGCAATTTAC 1709
Db 1321 TGGTGGCTGCTGCTCCAGTGATTTTCAGTGCTCTATCAAGAGAGAGTAGCAATTTAC 1380
Qy 1710 CAGTGGTGATGGGAGTTCTTGGCCGGCATGGATCTTAATATCCAAAGTGTGATGAAGTCAG 1769
Db 1381 CAGTGGTGATGGGAGTTCTTGGCCGGCATGGATCTTAATATCCAAAGTGTGATGAAGTCAG 1440

QY 1770 AAGCTGGTATATTTTGAAGGCCCAAGACTCCCTTTAGAGCATCACCTGTACGTAGT 1829
Db 1441 AAGCTGGTATATTTTGAAGGCCCAAGACTCCCTTTAGAGCATCACCTGTACGTAGT 1500
QY 1830 CAGTTACGTAAATCCTGGAGAGTGCACAGGCTGACTGACCGTGGCTACTCACATTCTTG 1889
Db 1501 CAGTTACGTAAATCCTGGAGAGTGCACAGGCTGACTGACCGTGGCTACTCACATTCTTG 1560
QY 1890 CTGCATCAGTCAGCACTGTGACTTCTTTATTAAGTAAGTATAGTAAACCAAGAGAATCCACA 1949
Db 1561 CTGCATCAGTCAGCACTGTGACTTCTTTATTAAGTAAGTATAGTAAACCAAGAGAATCCACA 1620
QY 1950 CTGTGTGTCCTTTTACAAGCTATCAAGTCTGAAGTACCCAACTTGCAGAAAACAAAGGA 2009
Db 1621 CTGTGTGTCCTTTTACAAGCTATCAAGTCTGAAGTACCCAACTTGCAGAAAACAAAGGA 1680
QY 2010 ATTTTGGGCCACCAATTTTGGATTCAGCAGGTCTCTTCTGACTATATCTCTCCAGAAAT 2069
Db 1681 ATTTTGGGCCACCAATTTTGGATTCAGCAGGTCTCTTCTGACTATATCTCTCCAGAAAT 1740
QY 2070 TTTCTCTTTTGAAGTACTACTGGATTTACATTTGATGGATGCTCTTACAAGCCTCATGA 2129
Db 1741 TTTCTCTTTTGAAGTACTACTGGATTTACATTTGATGGATGCTCTTACAAGCCTCATGA 1800
QY 2130 TCTACAGCTGGAAAGAAATATCTACTGTGCTGTTTCAATATGATGGTGTCTCTCAGGTGCA 2189
Db 1801 TCTACAGCTGGAAAGAAATATCTACTGTGCTGTTTCAATATGATGGTGTCTCTCAGGTGCA 1860
QY 2190 GTTGGTGAAATACGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCTTAGCCTCTCT 2249
Db 1861 GTTGGTGAAATACGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCTTAGCCTCTCT 1920
QY 2250 AdGTTATGTGTTGTAGTATAGACAAACAGGGGATCCTGTCAACAGAGGCTTAAATTGA 2309
Db 1921 AGTTATGTGTTGTAGTATAGACAAACAGGGGATCCTGTCAACAGAGGCTTAAATTGA 1980
QY 2310 AdGGCCCTTTAAATATATAATATGAGTCAAAATAGAAATGACATCAGGTGGGAAGACTCCA 2369
Db 1981 AdGGCCCTTTAAATATATAATATGAGTCAAAATAGAAATGACATCAGGTGGGAAGACTCCA 2040
QY 2370 ATATCTAGCTTCTCGATATGATTTCAATTGACTTAGATCGTGTGGGCATCCACGGCTGGTC 2429
Db 2041 ATATCTAGCTTCTCGATATGATTTCAATTGACTTAGATCGTGTGGGCATCCACGGCTGGTC 2100
QY 2430 CTATGGAGGATACCTCTCCCTGATGGCATTAAATGACAGGTCAGATATCTTTCAGGGTTGC 2489
Db 2101 CTATGGAGGATACCTCTCCCTGATGGCATTAAATGACAGGTCAGATATCTTTCAGGGTTGC 2160
QY 2490 TATTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAAACGTTA 2549
Db 2161 TATTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAAACGTTA 2220
QY 2550 TATGGGTACCTGACAGAAATGAAACGGGCTTACTTAGATCTGTGGCCATGCAAGC 2609
Db 2221 TATGGGTACCTGACAGAAATGAAACGGGCTTACTTAGATCTGTGGCCATGCAAGC 2280
QY 2610 AGAAAGTTCCTCTGACCAATCTTTTACTGCTCTTACATGTTTCTCTGGATGAGAA 2669
Db 2281 AGAAAGTTCCTCTGACCAATCTTTTACTGCTCTTACATGTTTCTCTGGATGAGAA 2340
QY 2670 TGTCCATTTTGCACATACCAAGTATATTAATCTAGTGTGTTTGTAGTGGGCTGGAAGGCCATA 2729
Db 2341 TGTCCATTTTGCACATACCAAGTATATTAATCTAGTGTGTTTGTAGTGGGCTGGAAGGCCATA 2400
QY 2730 TGAATTACAGATCTATCTCAGGAGAGACACAGATTAAGATTCCTGATCGGAGAACCA 2789
Db 2401 TGAATTACAGATCTATCTCAGGAGAGACACAGATTAAGATTCCTGATCGGAGAACCA 2460
QY 2790 TTATGCACTGATCTTTTGGCACTTACCTTCAAGAAAACCTTGGATCACGTTATTTGCTGCT 2849
Db 2461 TTATGCACTGATCTTTTGGCACTTACCTTCAAGAAAACCTTGGATCACGTTATTTGCTGCT 2520
QY 2850 AAAAGTGATATAATTTTGAACCTGTGTAGAACTCTCTGGTATATACACTGGCTATTTAAACCA 2909

Db 2521 AAAAGTGATATAATTTTGAACCTGTGTAGAACTCTCTGGTATACACTGGCTATTTAAACCA 2580
QY 2910 ATGAGGAGGTTTAAATCAACAGAAAACACAGAAATTGATCATCATCTTTTGATACCTGCCAT 2969
Db 2581 ATGAGGAGGTTTAAATCAACAGAAAACACAGAAATTGATCATCATCTTTTGATACCTGCCAT 2640
QY 2970 GTAACTACTCTCTCTGAAAATAAATGTGTGCCATGACGGGCTCTACGGTTTGTGGTAGT 3029
Db 2641 GTAACTACTCTCTCTGAAAATAAATGTGTGCCATGACGGGCTCTACGGTTTGTGGTAGT 2700
QY 3030 AATCTAATACCTTTAAACCCACATGCTCAAAATCAATCAATGATACATATCTCTGAGAGACCA 3089
Db 2701 AATCTAATACCTTTAAACCCACATGCTCAAAATCAATCAATGATACATATCTCTGAGAGACCA 2760
QY 3090 GCAATACCATTAAGAAATTTACTTAAAAA 3120
Db 2761 GCAATACCATTAAGAAATTTACTTAAAAA 2791
RESULT 11
ACA92425
ID ACA92425 standard; DNA; 2929 BP.
XX
AC ACA92425;
XX
DT 15-JUL-2003 (first entry)
XX
DE DNA encoding human PMM-10.
XX
KW Human; protein modification and maintenance molecule; PMM; cancer;
KW cell proliferation disorder; atherosclerosis; neurological disorder;
KW epilepsy; Huntington's disease; stroke; immune disorder; allergy;
KW inflammatory disorder; AIDS; developmental disorder; hypothyroidism;
KW Cushing's syndrome; gastrointestinal disorder; epithelial disorder;
KW infection; cystostatic; antiatherosclerotic; anticonvulsant; nootropic;
KW neuroprotective; cerebroprotective; anti-HIV; antiallergic; vulnerary;
KW antiinflammatory; thymimetic; gene; ds.
OS Homo sapiens.
XX
PN WO2003031939-A2.
XX
PD 17-APR-2003.
XX
PF 11-OCT-2002; 2002WO-US032850.
XX
PR 12-OCT-2001; 2001US-0329689P.
PR 25-OCT-2001; 2001US-0335703P.
PR 09-NOV-2001; 2001US-034887P.
PR 28-NOV-2001; 2001US-034145P.
PR 06-DEC-2001; 2001US-0337451P.
PR 14-DEC-2001; 2001US-0340584P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Ramkumar J, Gorvad AE, Baughn MR, Emerling BM, Yang J, Lee SY;
PI Tran UX, Becha SD, Duggan BM, Lee BA, Griffin JA, Li JX;
PI Sprague WW, Hafalia AJA, Chawla NK, Lehr-Mason PM, Kable AE, Yue H;
PI Marquis JP, Yao MG, Richardson TW, Tang TY, Jin P, Chien D;
PI Bhatia U, Burrill JD, Lee S, Blake JJ, Ho A, Zheng W;
XX
DR WPI: 2003-430274/40.
DR P-PSDB; ABU92030.
XX
PT New human protein modification and maintenance molecules (PMM), useful
PT for diagnosing; treating and preventing diseases or conditions associated
PT with the aberrant PMM expression e.g. cancer, atherosclerosis, or
PT infections.
XX
PS Claim 5; Page 289; 311pp; English.
XX
CC The present invention relates to the isolation of human protein

Db	2374	CQCTGACCAGAAATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAAAGTT	2433
Qy	2619	CCCTCTGAAACCAATCGTTTACTGCTTTCATCGTGGTTTCTCGATGAGAAATGTCATTT	2678
Db	2434	CCCTCTGAAACCAATCGTTTACTGCTTTCATCGTGGTTTCTCGATGAGAAATGTCATTT	2493
Qy	2679	TGCACATACCACTATATTACTGAGTTTATTTAGTCAGGGCTGGAAGCCATATGATTTACA	2738
Db	2494	TGCACATACCACTATATTACTGAGTTTATTTAGTCAGGGCTGGAAGCCATATGATTTACA	2553
Qy	2739	GATCTATCTCTCAGGAGACACAGCATAAAGATTTCTGAAATCGGAGAACATATGAAT	2798
Db	2554	GATCTATCTCTCAGGAGACACAGCATAAAGATTTCTGAAATCGGAGAACATATGAAT	2613
Qy	2799	GCATCTTTTGCACTACTTCAAGAAAACCTTTGGATCAGGTATTCCTGCTCTAAAGTGAT	2858
Db	2614	GCATCTTTTGCACTACTTCAAGAAAACCTTTGGATCAGGTATTCCTGCTCTAAAGTGAT	2673
Qy	2859	ATAATTTTGACCTGTGTGAACCTCTCTGTATACACTGGCTATTAAACCAATGAGGAG	2918
Db	2674	ATAATTTTGACCTGTGTGAACCTCTCTGTATACACTGGCTATTAAACCAATGAGGAG	2733
Qy	2919	TTTAATCAACAGAAAACACAGAAATTTGATCATCATATTTGATACCTGCCCATGTAATCT	2978
Db	2734	TTTAATCAACAGAAAACACAGAAATTTGATCATCATATTTGATACCTGCCCATGTAATCT	2793
Qy	2979	ATCTCTGAAATTAATTTGGTGGCCATGAGGGTCTACGGTTTGTGGTAGTAATCTAATA	3038
Db	2794	ATCTCTGAAATTAATTTGGTGGCCATGAGGGTCTACGGTTTGTGGTAGTAATCTAATA	2853
Qy	3039	CTTTAAACCCACATGCTCAAAATCAAAATGATACATATTCCTGAGAGACCCAGCAATACCA	3098
Db	2854	CTTTAAACCCACATGCTCAAAATCAAAATGATACATATTCCTGAGAGACCCAGCAATACCA	2913
Qy	3099	TAAGAAATTAATAAAA 3114	
Db	2914	TAAGAAATTAATAAAA 2929	
RESULT 12			
ACA92421			
ID	ACA92421 standard; DNA; 2952 BP.		
XX	ACA92421;		
AC	ACA92421;		
DT	15-JUL-2003 (first entry)		
XX	DNA encoding human PMM-6.		
KW	Human; protein modification and maintenance molecule; PMM; cancer;		
KW	cell proliferation disorder; atherosclerosis; neurological disorder;		
KW	epilepsy; Huntington's disease; stroke; immune disorder; allergy;		
KW	inflammatory disorder; AIDS; developmental disorder; hypothyroidism;		
KW	Cushing's syndrome; gastrointestinal disorder; epithelial disorder;		
KW	infection; cytostatic; antiarteriosclerotic; anticonvulsant; nootropic;		
KW	neuroprotective; cerebroprotective; anti-HIV; antiallergic; vulnerary;		
KW	antiinflammatory; thylomimetic; gene; ds.		
OS	Homo sapiens.		
XX	WO2003031939-A2.		
PN	17-APR-2003.		
XX	11-OCT-2002; 2002WO-US032850.		
XX	12-OCT-2001; 2001US-0329689P.		
PR	25-OCT-2001; 2001US-0335703P.		
PR	09-NOV-2001; 2001US-0348887P.		
PR	28-NOV-2001; 2001US-0334145P.		
PR	06-DEC-2001; 2001US-0337451P.		
PR	14-DEC-2001; 2001US-0340584P.		

PA	(INCY-) INCYTE GENOMICS INC.	
XX	Ramkumar J, Gorvad AE, Baughn MR, Emerling BM, Yang J, Lee SY;	
PI	Tran UK, Becha SD, Duggan BM, Lee EA, Griffin JA, Li JX;	
PI	Sprague WW, Hafalia AJA, Chawla NK, Lehr-Mason PM, Kable AE, Yue H;	
PI	Marquis JP, Yao MG, Richardson TW, Tang TY, Jin P, Chien D;	
PI	Bhatia U, Burrill JD, Lee S, Blake JJ, Ho A, Zheng W;	
XX	WPI; 2003-430274/40.	
DR	P-PSDB; ABU92026.	
XX	New human protein modification and maintenance molecules (PMM), useful	
PT	for diagnosing, treating and preventing diseases or conditions associated	
PT	with the aberrant PMM expression e.g. cancer, atherosclerosis, or	
PT	infections.	
PS	Claim 5; Page 285-286; 311pp; English!	
XX	The present invention relates to the isolation of human protein	
CC	modification and maintenance molecules (PMM), and the polynucleotide	
CC	sequences encoding them. A total of 40 PMM polypeptides (designated PMM	
CC	-1 to PMM-40) are disclosed. The sequences of the invention are useful	
CC	for diagnosing a condition or disease associated with the expression of	
CC	PMM in a subject, preparing a polyclonal or monoclonal antibody, and	
CC	generating an expression profile of a sample containing the	
CC	polynucleotides. The diseases or conditions associated with decreased	
CC	expression or overexpression of PMM are cell proliferation disorders	
CC	(e.g. cancer, atherosclerosis), neurological disorders (e.g. epilepsy,	
CC	Huntington's disease, stroke), immune/inflammatory disorders, (e.g. AIDS,	
CC	allergies), developmental disorders (e.g. hypothyroidism, Cushing's	
CC	syndrome), gastrointestinal or epithelial disorders, and infections. The	
CC	PMM polypeptides or their fragments are useful in screening compounds	
CC	for effectiveness as agonists or antagonists of the polypeptides, or in	
CC	altering the expression of the target polynucleotide and compounds that	
CC	specifically bind to, or modulate the activity of the polypeptide.	
CC	ACA92416-ACA92455 encode the human PMM polypeptides of the invention	
XX	Sequence 2952 BP; 886 A; 604 C; 664 G; 798 T; 0 U; 0 Other;	
Query Match		72.8%; Score 2271; DB 10; Length 2952;
Best Local Similarity		99.9%; Pred. No! 0;
Matches 2321; Conservative		0; Mismatches 1; Indels 0; Gaps 0;
Qy	11	GCCTCCGAGGCCAAGCCCGCTCTACTGCGCGCGTCTTCTTAGTCCGCGTTCGCGGC 70
Db	1	GCCTCCGAGGCCAAGCCCGCTCTACTGCGCGCGTCTTCTTAGTCCGCGTTCGCGGC 60
Qy	71	CTGGTTTGTACCG 130
Db	61	CTGGTTTGTACCG 120
Qy	131	GGCGCAGCATGAAGCG 190
Db	121	GGCGCAGCATGAAGCG 180
Qy	191	GCCGGGGGGAAGAAATGCAATGCGAGCGAATGCAACAGCAACAGCAACAGCGTCTTG 250
Db	181	GCCGGGGGGAAGAAATGCAATGCGAGCGAATGCAACAGCAACAGCAACAGCGTCTTG 240
Qy	251	AGATATTTGAACTGCGGCGCTGTGAGGAGAAATTTGAATCAGGATCGGCTTAAATGG 310
Db	241	AGATATTTGAACTGCGGCGCTGTGAGGAGAAATTTGAATCAGGATCGGCTTAAATGG 300
Qy	311	AGCCTTTTATTTAGCGCGGTATTTCTCTGGAGTCAGCTTAAAGCTGTTGCCGATACCA 370
Db	301	AGCCTTTTATTTAGCGCGGTATTTCTCTGGAGTCAGCTTAAAGCTGTTGCCGATACCA 360
Qy	371	GAAATATCATGCTACATGCTGAAGCCACCATGATTTTCATGTTGTGAAGAGA 430
Db	361	GAAATATCATGCTACATGCTGAAGCCACCATGATTTTCATGTTGTGAAGAGA 420
Qy	431	ATGATCCAGATGACCTTATTCAGACAGATCTATTACCTTCGCCATGCTCTGTGTGAACA 490

Db 421 ATGATCCAGATGGACCTCAITTCAGACAGAATCTATTACCTTCCCATGTCTGTGTGAGACA 480
Qy 491 GAGAAATACACTGTTTATTCTGAAATCCCAAACTATCAATAGACAGCAGCTCTTAA 550
Db 481 GAGAAATACACTGTTTATTCTGAAATCCCAAACTATCAATAGACAGCAGCTCTTAA 540
Qy 551 TGCTCTCTTGGAGCCCTCTTTTGGATCTTTTTCAGGCAACACTGAGCTATGGAATGATT 610
Db 541 TGCTCTCTTGGAGCCCTCTTTTGGATCTTTTTCAGGCAACACTGAGCTATGGAATGATT 600
Qy 611 CTCGAGAGAGAGACTATTAAAGAGAAAGAAACGCATTTGGAACAGTCGGAAATTCGCTCTT 670
Db 601 CTCGAGAGAGAGACTATTAAAGAGAAAGAAACGCATTTGGAACAGTCGGAAATTCGCTCTT 660
Qy 671 ACGATTATCCCAAGAGAGTGGAAATTTCTGTTTCAAGCCGGTAGTGGAAATTTATCAAG 730
Db 661 ACGATTATCCCAAGAGAGTGGAAATTTCTGTTTCAAGCCGGTAGTGGAAATTTATCAAG 720
Qy 731 TAAAGATGAGAGGCCCAAGAGATTTACGAAACAACTTTTAAAGCCCAATCTAGTGGAA 790
Db 721 TAAAGATGAGAGGCCCAAGAGATTTACGAAACAACTTTTAAAGCCCAATCTAGTGGAA 780
Qy 791 CTAGTTGTCCCAACATACGATGGATTTGGAATCTTAAATTTATGCCCGCTGATCCAGACTGGATTG 850
Db 781 CTAGTTGTCCCAACATACGATGGATTTGGAATCTTAAATTTATGCCCGCTGATCCAGACTGGATTG 840
Qy 851 CTTTATATACATACGATGGATTTGGAATCTTAAATTTATGCCCGCTGATCCAGACTGGAGAC 910
Db 841 CTTTATATACATACGATGGATTTGGAATCTTAAATTTATGCCCGCTGATCCAGACTGGAGAC 900
Qy 911 TCACCTATGTGCAATGAGCTAGCCAACTGGAAGAGATGCCAGATCAGCTGGAGTCG 970
Db 901 TCACCTATGTGCAATGAGCTAGCCAACTGGAAGAGATGCCAGATCAGCTGGAGTCG 960
Qy 971 CTACCTTTGTTCTCCAAAGAAATTTGATAGATTTCTGGCTATTTGGTGTGTCCTAAAG 1030
Db 961 CTACCTTTGTTCTCCAAAGAAATTTGATAGATTTCTGGCTATTTGGTGTGTCCTAAAG 1020
Qy 1031 CTGAAACAACTCCAGTGTGTGTAATTTCTAGATTTCTATATGAGAAATGATGAT 1090
Db 1021 CTGAAACAACTCCAGTGTGTGTAATTTCTAGATTTCTATATGAGAAATGATGAT 1080
Qy 1091 CTGAGTGGAAATTTATTCATGTTACATCCCTATTTGGAAACAGAGGGGCGAGTTCAAT 1150
Db 1081 CTGAGTGGAAATTTATTCATGTTACATCCCTATTTGGAAACAGAGGGGCGAGTTCAAT 1140
Qy 1151 TCCGTTATCTTAAACAGGTACAGCAATCTTAAAGTCACTTTTAAAGATGTGAGAAATAA 1210
Db 1141 TCCGTTATCTTAAACAGGTACAGCAATCTTAAAGTCACTTTTAAAGATGTGAGAAATAA 1200
Qy 1211 TGAATGATCTGAGAGAGGATCATAGATGTATAGATAAGGAACCTAAATTCACCTTTTG 1270
Db 1201 TGAATGATCTGAGAGAGGATCATAGATGTATAGATAAGGAACCTAAATTCACCTTTTG 1260
Qy 1271 AGATTCATTTTGAAGAGTTGAATATATGCGAGAGCTGGATGGACTCTCTGAGGGAATAA 1330
Db 1261 AGATTCATTTTGAAGAGTTGAATATATGCGAGAGCTGGATGGACTCTCTGAGGGAATAA 1320
Qy 1331 ATGCTTGGTCCATCTTACTAGATCGCTCCAGACTCGCCCTACAGATAGTGTGATCTCAC 1390
Db 1321 ATGCTTGGTCCATCTTACTAGATCGCTCCAGACTCGCCCTACAGATAGTGTGATCTCAC 1380
Qy 1391 CTGAATTTATTTTCCAGTGAAGATGATTTTGAAGGCGAGAGCTCAATTTAGTGTGAGT 1450
Db 1381 CTGAATTTATTTTCCAGTGAAGATGATTTTGAAGGCGAGAGCTCAATTTAGTGTGAGT 1440
Qy 1451 TGCCCTGATTTCTGTAGCGCCTAATTTATCTATGAAGAAACACAGACATCTGGATAATA 1510
Db 1441 TGCCCTGATTTCTGTAGCGCCTAATTTATCTATGAAGAAACACAGACATCTGGATAATA 1500
Qy 1511 TCCATGACATCTTTTCATGTTTTCCTCCAAAGTCCAGAGGAAATTCAGTTATTTTGG 1570
Db 1501 TCCATGACATCTTTTCATGTTTTCCTCCAAAGTCCAGAGGAAATTCAGTTATTTTGG 1560

Qy 1571 CCTCTGAATGCAAAAACAGGTTTCGGTCATTTATACAAAATTACATCTATTTTAAAGGAAA 1630
Db 1561 CCTCTGAATGCAAAAACAGGTTTCGGTCATTTATACAAAATTACATCTATTTTAAAGGAAA 1620
Qy 1631 GCAAAATATAAACAAGTCCAGTGGTGGCTGCTCTCCAAAGTGAATTTCAAGTGTCTCTATCA 1690
Db 1621 GCAAAATATAAACAAGTCCAGTGGTGGCTGCTCTCCAAAGTGAATTTCAAGTGTCTCTATCA 1680
Qy 1691 AAGAGGAGATAGCAAAATTTACAGTGGTGAATGGGAAAGTTCTTGGCCGGCATGGATCTAATA 1750
Db 1681 AAGAGGAGATAGCAAAATTTACAGTGGTGAATGGGAAAGTTCTTGGCCGGCATGGATCTAATA 1740
Qy 1751 TCCAAGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGGCCAACAAAGACTCCCTTTTAG 1810
Db 1741 TCCAAGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGGCCAACAAAGACTCCCTTTTAG 1800
Qy 1811 AGCATACCTGTACGTAGTCAGTTAGTAAATCTCGAGAGGTGACAGGCTGACTGACC 1870
Db 1801 AGCATACCTGTACGTAGTCAGTTAGTAAATCTCGAGAGGTGACAGGCTGACTGACC 1860
Qy 1871 GTGGCTACTCACATTTCTGCTGCATCAGTCAGTCACTGCTCTTTTATAGTAAAGTATA 1930
Db 1861 GTGGCTACTCACATTTCTGCTGCATCAGTCAGTCACTGCTCTTTTATAGTAAAGTATA 1920
Qy 1931 GTAACCCAGAGAGATCCACACTGTGTCTGCTTTTACAGCTATCAAGTCTCTGAAGATGACC 1990
Db 1921 GTAACCCAGAGAGATCCACACTGTGTCTGCTTTTACAGCTATCAAGTCTCTGAAGATGACC 1980
Qy 1991 CAACCTTGCAAAAACAAAGGAATTTTGGGCCACCATTTTGGATTCAGAGGTCCTCTTCCTG 2050
Db 1981 CAACCTTGCAAAAACAAAGGAATTTTGGGCCACCATTTTGGATTCAGAGGTCCTCTTCCTG 2040
Qy 2051 ACTATCTCCTCCAGAAATTTTCTCTTTGAAAGTACTTCTGATTTTACATGTTATGGGA 2110
Db 2041 ACTATCTCCTCCAGAAATTTTCTCTTTGAAAGTACTTCTGATTTTACATGTTATGGGA 2100
Qy 2111 TGCTCTPACAGCCTCATGATCTACAGCCTGGAAGAAATATCTCTACTGTCTGCTTCATAT 2170
Db 2101 TGCTCTPACAGCCTCATGATCTACAGCCTGGAAGAAATATCTCTACTGTCTGCTTCATAT 2160
Qy 2171 ATGGTGTCTCCTCAGGTGAGTTGGTGAATATCGGTTTAAAGGAGTCAAGTATTTCCGCT 2230
Db 2161 ATGGTGTCTCCTCAGGTGAGTTGGTGAATATCGGTTTAAAGGAGTCAAGTATTTCCGCT 2220
Qy 2231 TGAATACCTTAGCCTCTCTAGGTTATGTTGTTAGTATAGACAAACAGGGGATCTGTC 2290
Db 2221 TGAATACCTTAGCCTCTCTAGGTTATGTTGTTAGTATAGACAAACAGGGGATCTGTC 2280
Qy 2291 ACCGAGGGCTTAAATTTTGAAGGCCCTTTAAATATAAAATGG 2332
Db 2281 ACCGAGGGCTTAAATTTTGAAGGCCCTTTAAATATAAAATGG 2322

RESULT 13

AD116386

ID AD116386 standard; DNA; 2696 BP.

XX

AC AD116386;

XX

DT 22-APR-2004 (first entry)

DX

Human protein modification and maintenance molecule (PMM) gene #19.

DE

DE

KW

KW

KW

KW

KW

KW

KW

KW

OS

human; protein modification and maintenance molecule; PMM;
Gastrointestinal disorders; peptic ulcer; Crohn's disease;
cardiovascular disorders; hypertension; congenital heart disease;
autoimmune disease; inflammatory disease; AIDS; anaemia;
developmental disorder; Cushing's syndrome; tubular acidosis;
epithelial disorder; eczema; scabies; neurological disorder;
Alzheimer's disease; multiple sclerosis; infection; cancer; ds.

QY 1513 CATGACATCTTTTCATGTTTCCCAAGTCACGAGAGGAATTCAGTTATTTTGGC 1572
Db 1502 CATGACATCTTTTCATGTTTCCCAAGTCACGAGAGGAATTCAGTTATTTTGGC 1561
QY 1573 TCTGAATGCAAAACAGGTTTCCTGATTTATACAAATTTACATCTATTTTAAAGGAAGC 1632
Db 1562 TCTGATGCAAAACAGGTTTCCTGATTTATACAAATTTACATCTATTTTAAAGGAAGC 1621
QY 1633 AAATATAAACGATCCAGTGTGGGCTCCCTGCTCCAAAGTATTCAAAGTTCCTATCAAA 1692
Db 1622 AAATATAAACGATCCAGTGTGGGCTCCCTGCTCCAAAGTATTCAAAGTTCCTATCAAA 1681
QY 1693 GAGGAGATAGCAATACACATGTGTGAATGGAGTTCTTGGCCGATGATCTTAATATC 1752
Db 1682 GAGGAGATAGCAATACACATGTGTGAATGGAGTTCTTGGCCGATGATCTTAATATC 1741
QY 1753 CAAGTTGATGAAGTCAGAAGGCTGGTATATTTTGAAGGCACCAAGACTCCCTTTAGAG 1812
Db 1742 CAAGTTGATGAAGTCAGAAGGCTGGTATATTTTGAAGGCACCAAGACTCCCTTTAGAG 1801
QY 1813 CATCACTGTACGTAGTCAGTTACGTAAATCTCTGGAGAGGTGACAAAGCTGACTGACCGT 1872
Db 1802 CATCACTGTACGTAGTCAGTTACGTAAATCTCTGGAGAGGTGACAAAGCTGACTGACCGT 1861
QY 1873 GGCTACTCACATCTTCTGCTCATCAGTCAGCACTGTGACTTCTTTATAAGTAACTATAGT 1932
Db 1862 GGCTACTCACATCTTCTGCTCATCAGTCAGCACTGTGACTTCTTTATAAGTAACTATAGT 1921
QY 1933 AACCAAGAAGATCCACATGTGTGTCCTTTACAGCTATCAAGTCTCTGAGATGACCCA 1992
Db 1922 AACCAAGAAGATCCACATGTGTGTCCTTTACAGCTATCAAGTCTCTGAGATGACCCA 1981
QY 1993 ACTTGCAAAACAAAGGAATTTGGGCCACCAATTTGGATTCAGCAGTCTCTTCTTGAC 2052
Db 1982 ACTTGCAAAACAAAGGAATTTGGGCCACCAATTTGGATTCAGCAGTCTCTTCTTGAC 2041
QY 2053 TATACTCTCCAGAAATTTCTTTTGAAGTACTACTCGAATTAATGATGAGGATG 2112
Db 2042 TATACTCTCCAGAAATTTCTTTTGAAGTACTACTCGAATTAATGATGAGGATG 2101
QY 2113 CTCTACAGCCTCATGATCTACAGCTGGAAGAAATATCTACTGTGCTTCAATAT 2172
Db 2102 CTCTACAGCCTCATGATCTACAGCTGGAAGAAATATCTACTGTGCTTCAATAT 2161
QY 2173 GGTGGTCTCAGGTGAGTGTGTAATATCGGTTTAAAGGAGTCAAGTATTTCCGCTTG 2232
Db 2162 GGTGGTCTCAGGTGAGTGTGTAATATCGGTTTAAAGGAGTCAAGTATTTCCGCTTG 2221
QY 2233 AATACCTAGCCTCTCTAGGTTATGTGGTTGTAGTGATAGACACAGGGGATCCTGTAC 2292
Db 2222 AATACCTAGCCTCTCTAGGTTATGTGGTTGTAGTGATAGACACAGGGGATCCTGTAC 2281
QY 2293 CGAGGGCTTAAATTTGAAGGCGCTTTAAATATATAAATG 2331
Db 2282 CGAGGGCTTAAATTTGAAGGCGCTTTAAATATATAAATG 2320

RESULT 14
ABK83325
ID ABK83325 standard; cDNA; 4523 BP.
XX AC ABK83325;
XX AC
XX AC
DT 12-AUG-2002 (first entry)
DE cDNA encoding human DPRP-1 splice variant #1.
XX
KW Human; serine protease; dipeptidyl peptidase IV-related protein; DPRP;
KW DPPIV; infection; human immunodeficiency virus; HIV-1; HIV-2; pain;
KW diabetes; infertility; obesity; anorexia; Parkinson's disease; stroke;
KW heart failure; hypertension; urinary retention; osteoporosis; cancer;
KW ulcer; allergy; cancer; psychotic disorder; neurological disorder;

KW dyskinesia; reproductive disorder; inflammatory disorder;
KW metabolic disorder; gene; ss.
OS Homo sapiens.
XX WO200231134-A2.
XX 18-APR-2002.
XX 12-OCT-2001; 2001WO-US031874.
XX 12-OCT-2000; 2000US-0240117P.
XX (FERR) FERRING BV.
XX Qi S, Akinsanya KO, Riviere PJ, Junien J;
XX WPI; 2002-444178/47.
XX P-PSDB; ABG61594.
XX New dipeptidyl peptidase IV-related proteins and nucleic acids encoding
XX the proteins, useful for treating e.g. fungal, bacterial, protozoan and
XX viral infections, cancers, allergies, neurological disorders, or pain.
XX Disclosure; Page 61-62; 113pp; English.
XX The present invention relates to the isolation of novel human serine
XX proteases referred to as dipeptidyl peptidase IV (DPPIV)-related proteins
XX (DPRP). The dipeptidyl peptidase IV-related proteins (DPRP) and nucleic
XX acids encoding them are useful for treating infections such as fungal,
XX bacterial, protozoan and viral infections, particularly infections caused
XX by human immunodeficiency virus (HIV-1 or HIV-2), pain, diabetes,
XX precocious puberty, infertility, obesity, anorexia, bulimia, Parkinson's
XX disease, acute heart failure, hypotension, hypertension, urinary
XX retention, osteoporosis, angina pectoris, stroke, ulcers, asthma,
XX allergies, cancers, migraines, vomiting, psychotic and neurological
XX disorders (e.g. anxiety, dementia, or schizophrenia), and dyskinesias.
XX These may also be used in discovering therapeutic agents for the
XX treatment of reproductive, inflammatory and metabolic disorders. ABK83322
XX -ABK83343 encode human DPRP proteins
SQ Sequence 4523 BP; 1384 A; 828 C; 940 G; 1371 T; 0 U; 0 Other;
Query Match 68.2%; Score 2128; DB 6; Length 4523;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2178; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 AAGTGCTAAAGCTCTCGAGGCCCAAGGCCGCTGCTACTGCGCCGCTCTCTTAGTGCCG 60
Db 1 AAGTGCTAAAGCTCTCGAGGCCCAAGGCCGCTGCTACTGCGCCGCTCTCTTAGTGCCG 60
QY 61 CGTTCCCGCTGGGTTGTACCGCGCCGCGCGGAGGAGCCACTGCAACCAAGACCG 120
Db 61 CGTTCCCGCTGGGTTGTACCGCGCCGCGCGGAGGAGCCACTGCAACCAAGACCG 120
QY 121 GAGTGAGGGGGCGGCGGAGCATGAGCGGCGAGGCGCGCTCTCATAGCCAGCTGGGACCG 180
Db 121 GAGTGAGGGGGCGGCGGAGCATGAGCGGCGAGGCGCGCTCTCATAGCCAGCTGGGACCG 180
QY 181 TCCGGCGGGGGCGGGGGAGGAAATGCAACATGCGCAGCAGCAATGGAACACAGACAG 240
Db 181 TCCGGCGGGGGCGGGGGAGGAAATGCAACATGCGCAGCAGCAATGGAACACAGACAG 240
QY 241 CTGGGTTGAGATATTTGAACTCGGACTGTGAGGAGATATTTGAATCAGAGGATCGG 300
Db 241 CTGGGTTGAGATATTTGAACTCGGACTGTGAGGAGATATTTGAATCAGAGGATCGG 300
QY 301 CCTAAATGGAGCCCTTTTATGTTGAGCGGTATTTCTTGGAGTCTGAGGAGTCTT 360
Db 301 CCTAAATGGAGCCCTTTTATGTTGAGCGGTATTTCTTGGAGTCTGAGGAGTCTT 360
QY 361 GCGGATACAGAAAAATATCATGGCTTACATGATGCTAAGGCACCATGATTTTCATGTTT 420

```
Db 361 GCCATACCAGAAAATATCATGGCTACATGATGGCTAAGGCCACCACATGATTTTCATGTTT 420
QY 421 GTGAAGAGGAATGATCCAGATGACCTCATTGAGACAGAACTATTATACCTTTGCCATGCT 480
Db 421 GTGAAGAGGAATGATCCAGATGACCTCATTGAGACAGAACTATTATACCTTTGCCATGCT 480
QY 481 GGTGAGAACAGAGAAAATACACTGTTTATTCTGAAATTCCTGAAATTCCTCAATAGAGCA 540
Db 481 GGTGAGAACAGAGAAAATACACTGTTTATTCTGAAATTCCTCAATAGAGCA 540
QY 541 GCACTCTTAATGCTCTCTGGAAGCCTCTTTTGGATCTTTTTCAGGCAACACTGGACTAT 600
Db 541 GCACTCTTAATGCTCTCTGGAAGCCTCTTTTGGATCTTTTTCAGGCAACACTGGACTAT 600
QY 601 GGAATGATTCTTCGAGAAGAGAACTATTAAAGAGAAAGAAAACGCAATTTGGAACAGTCGGA 660
Db 601 GGAATGATTCTTCGAGAAGAGAACTATTAAAGAGAAAGAAAACGCAATTTGGAACAGTCGGA 660
QY 661 ATTGCTTCTTACGATTATACCAAGGAAGTGGAACTTTCTGTTTCAAGCCGGTAGTGA 720
Db 661 ATTGCTTCTTACGATTATACCAAGGAAGTGGAACTTTCTGTTTCAAGCCGGTAGTGA 720
QY 721 ATTATACAGTAAAGAGTGGAGGCCACAGAGGATTTACGCAACACCTTTTAAGGCCAAT 780
Db 721 ATTATACAGTAAAGAGTGGAGGCCACAGAGGATTTACGCAACACCTTTTAAGGCCAAT 780
QY 781 CTAGTGGAACTAGTGTCTCCCAACATACGAGTGGATTCGCAACACCTTTTAAGGCCAAT 840
Db 781 CTAGTGGAACTAGTGTCTCCCAACATACGAGTGGATTCGCAACACCTTTTAAGGCCAAT 840
QY 841 CACTGGATTTGCTTTTATACATAGCAACGATATTGGATATCTAAACATCGTAACAGAGAA 900
Db 841 CACTGGATTTGCTTTTATACATAGCAACGATATTGGATATCTAAACATCGTAACAGAGAA 900
QY 901 GAAAGGAGACTCACTTATGTGCA CAATGAGCTAGGCCAACATGGAAGAGATGCCAGATCA 960
Db 901 GAAAGGAGACTCACTTATGTGCA CAATGAGCTAGGCCAACATGGAAGAGATGCCAGATCA 960
QY 961 GCTGGAGTCGCTACCTTTGTTCTCCAAAGAGAAATTTGATAGATATTTCTGGCTATTGGTGG 1020
Db 961 GCTGGAGTCGCTACCTTTGTTCTCCAAAGAGAAATTTGATAGATATTTCTGGCTATTGGTGG 1020
QY 1021 TGTCCAAAGCTGAAACAACTCCAGTGGTGGTAAATTTCTAGAAATTTCTATATGAAGAA 1080
Db 1021 TGTCCAAAGCTGAAACAACTCCAGTGGTGGTAAATTTCTAGAAATTTCTATATGAAGAA 1080
QY 1081 AATGATGAATCTGAGGTGGAAATTTATTCATGTTTACATCCCTTATGTTGGAACAGAGAGG 1140
Db 1081 AATGATGAATCTGAGGTGGAAATTTATTCATGTTTACATCCCTTATGTTGGAACAGAGAGG 1140
QY 1141 GAGATTCAATTCCTGTTATCCCTTAAACAGGTACAGCAATTCCTTAAAGTCACCTTTAAGATG 1200
Db 1141 GAGATTCAATTCCTGTTATCCCTTAAACAGGTACAGCAATTCCTTAAAGTCACCTTTAAGATG 1200
QY 1201 TCAGAAATPAATGATGATGCTGGAAGGAGATCATAGATGTCATAGATAAGGAACATAAT 1260
Db 1201 TCAGAAATPAATGATGATGCTGGAAGGAGATCATAGATGTCATAGATAAGGAACATAAT 1260
QY 1261 CAACCTTTTGAGATCTTATTGGAAGGATGGAATATTTGCCAGAGCTGGATGGACTCCT 1320
Db 1261 CAACCTTTTGAGATCTTATTGGAAGGATGGAATATTTGCCAGAGCTGGATGGACTCCT 1320
QY 1321 GAGGGAATATGCTGCTCCCTACTAGATCGCTCCAGACTCGCTACAGATAGTG 1380
Db 1321 GAGGGAATATGCTGCTCCCTACTAGATCGCTCCAGACTCGCTACAGATAGTG 1380
QY 1381 TTGATCTCACTGAAATTTATTCCTCAGTGAAGATGATGTTATGGAAGGACAGAGCTC 1440
Db 1381 TTGATCTCACTGAAATTTATTCCTCAGTGAAGATGATGTTATGGAAGGACAGAGCTC 1440
QY 1441 ATTGAGTCAGTGCCTGATTTCTGTGACGCCACTAAATTTATCTATGAAGAAACACAGACATC 1500
Db 1441 ATTGAGTCAGTGCCTGATTTCTGTGACGCCACTAAATTTATCTATGAAGAAACACAGACATC 1500
```

```
QY 1501 TGGATAAATATCATGACATCTTTTCATGTTTTCCTCCAAAGTCACGAGAGGAAATTTGAG 1560
Db 1501 TGGATAAATATCATGACATCTTTTCATGTTTTCCTCCAAAGTCACGAGAGGAAATTTGAG 1560
QY 1561 TTTATTTTTGCCCTCTGAATGCAAAACAGGTTTCCGTCATTTATATACAAATATACATCTATT 1620
Db 1561 TTTATTTTTGCCCTCTGAATGCAAAACAGGTTTCCGTCATTTATATACAAATATACATCTATT 1620
QY 1621 TTAAGGAAAGCAAAATATAAACGATCCAGTGGTGGGCTGCCTGCTCCAAAGTGAATTTCAAG 1680
Db 1621 TTAAGGAAAGCAAAATATAAACGATCCAGTGGTGGGCTGCCTGCTCCAAAGTGAATTTCAAG 1680
QY 1681 TGTCTTATCAAGAGGAGATAGCAATTTACAGTGGTGGTGAATGGGAAGTTCTTTGGCCGGCAT 1740
Db 1681 TGTCTTATCAAGAGGAGATAGCAATTTACAGTGGTGGTGAATGGGAAGTTCTTTGGCCGGCAT 1740
QY 1741 GGATCTAATATCCAAGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGGCCACCAAGAC 1800
Db 1741 GGATCTAATATCCAAGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGGCCACCAAGAC 1800
QY 1801 TCCCTTTTAGAGCATCACCTGTACGTAGTCAGTTACGTAAATCTCTGGAGAGGTGACAAGG 1860
Db 1801 TCCCTTTTAGAGCATCACCTGTACGTAGTCAGTTACGTAAATCTCTGGAGAGGTGACAAGG 1860
QY 1861 CTGACTGACCGTGGCTACTCACAATTTCTGCTGCATCAGTCAGTCACTGTAATTTTATA 1920
Db 1861 CTGACTGACCGTGGCTACTCACAATTTCTGCTGCATCAGTCAGTCACTGTAATTTTATA 1920
QY 1921 AGTAAGTATAGTAACAGAGAAATCCACACTGCTGTGCTCCCTTTACAAGCTATCAAGTCT 1980
Db 1921 AGTAAGTATAGTAACAGAGAAATCCACACTGCTGTGCTCCCTTTACAAGCTATCAAGTCT 1980
QY 1981 GAAGTGAACCCCACTTCCGCAAAACAAAGGAATTTTGGGCCACCATTTTGGATTCACAGGT 2040
Db 1981 GAAGTGAACCCCACTTCCGCAAAACAAAGGAATTTTGGGCCACCATTTTGGATTCACAGGT 2040
QY 2041 CCTCTTCTGACTATATCTCTCCAGAAAATTTCTCTTTTGAAGTACTACTGGATTTACA 2100
Db 2041 CCTCTTCTGACTATATCTCTCCAGAAAATTTCTCTTTTGAAGTACTACTGGATTTACA 2100
QY 2101 TTGTATGGGATGCTCTACAAGCCTCATGATCTACAGCCTGGAAAGAAATATCCTACTGTG 2160
Db 2101 TTGTATGGGATGCTCTACAAGCCTCATGATCTACAGCCTGGAAAGAAATATCCTACTGTG 2160
QY 2161 CTGTTTCATATATGTTGTC 2179
Db 2161 CTGTTTCATATATGTTGTC 2179
```

RESULT 15

ABK83331

ID ABK83331 standard; cDNA; 4676 BP.

XX AC ABK83331;

XX DT 12-AUG-2002 (first entry)

XX CD cDNA encoding human DPRP-1 splice variant #7.

Human; serine protease; dipeptidyl peptidase IV-related protein; DPRP; DPP4V; infection; human immunodeficiency virus; HIV-1; HIV-2; pain; diabetes; infertility; obesity; anorexia; Parkinson's disease; stroke; heart failure; hypertension; urinary retention; osteoporosis; cancer; ulcer; allergy; cancer; psychotic disorder; neurological disorder; dyskinesia; reproductive disorder; inflammatory disorder; metabolic disorder; gene; ss.

OS Homo sapiens.

XX WO200231134-A2.

XX PD 18-APR-2002.

XX 12-OCT-2001; 2001WO-US031874.
PF 12-OCT-2000; 2000US-0240117P.
PR (FERR) FERRING BV.
XX
XX Qi S, Akinsanya KO, Riviere PJ, Junien J;
XX WPI: 2002-444178/47.
XX P-P8DB; ABG61600.
DR
XX
XX New dipeptidyl peptidase IV-related proteins and nucleic acids encoding
PI the proteins, useful for treating e.g. fungal, bacterial, protozoan and
PT viral infections, cancers, allergies, neurological disorders, or pain.
XX
XX Disclosure; Page 72-73; 113pp; English.
XX
XX The present invention relates to the isolation of novel human serine
CC proteases referred to as dipeptidyl peptidase IV (DPPIV)-related proteins
CC (DPPP). The dipeptidyl peptidase IV-related proteins (DPPP) and nucleic
CC acids encoding them are useful for treating infections such as fungal,
CC bacterial, protozoan and viral infections, particularly infections caused
CC by human immunodeficiency virus (Hiv-1 or Hiv-2), pain, diabetes,
CC precocious puberty, infertility, obesity, anorexia, bulimia, Parkinson's
CC disease, acute heart failure, hypotension, hypertension, urinary
CC retention, osteoporosis, angina pectoris, stroke, ulcers, asthma,
CC allergies, cancers, migraine, vomiting, psychotic and neurological
CC disorders (e.g. anxiety, dementia, or schizophrenia), and dyskinesias.
CC These may also be used in discovering therapeutic agents for the
CC treatment of reproductive, inflammatory and metabolic disorders. ABK83322
CC -ABK83343 encode human DPPP proteins
XX
XX Sequence 4676 BP; 1424 A; 859 C; 979 G; 1414 T; 0 U; 0 Other;

Query Match 68.28; Score 2128; DB 6; Length 4676;
Best Local Similarity 99.98; Pred. No. 0;
Matches 2178; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAGTCTTAAGCCTCCAGGCCAAGCCGCTGCTACTGCGCGCGCTGCTTCTTAGTGCCG 60
DB 1 AAGTCTTAAGCCTCCAGGCCAAGCCGCTGCTACTGCGCGCGCTGCTTCTTAGTGCCG 60
QY 61 CGTTCCGCGCGCTGGGTTGTCAACCGCGCGCGCGCGAGGAGCCACTGCAACCAAGGACCG 120
DB 61 CGTTCCGCGCGCTGGGTTGTCAACCGCGCGCGCGCGAGGAGCCACTGCAACCAAGGACCG 120
QY 121 GAGTGGAGCGCGCGCGAGCATGAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
DB 121 GAGTGGAGCGCGCGCGAGCATGAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
QY 181 TCCGGCGCGCGCGCGGAGGAAATGCAACATGGCAGCAGCAATGGAAACAGAACAG 240
DB 181 TCCGGCGCGCGCGCGGAGGAAATGCAACATGGCAGCAGCAATGGAAACAGAACAG 240
QY 241 CTGGGTGTTGAGATATTTGAACTGCGGACTGTGAGGAGATATTTGAATCACAGGATCGG 300
DB 241 CTGGGTGTTGAGATATTTGAACTGCGGACTGTGAGGAGATATTTGAATCACAGGATCGG 300
QY 301 CCTAAATTTGAGCGCTTTTATGTTGAGCGGATTCCTGAGTCAGCTTAAAGAGCTGCTT 360
DB 301 CCTAAATTTGAGCGCTTTTATGTTGAGCGGATTCCTGAGTCAGCTTAAAGAGCTGCTT 360
QY 361 GCCGATACAGAAATATCATGCTACATGATGGCTTAGGCACCAATGATTTTCATGTTT 420
DB 361 GCCGATACAGAAATATCATGCTACATGATGGCTTAGGCACCAATGATTTTCATGTTT 420
QY 421 GTGAGAGGATCATCCAGATGCACTCATTCAGACAGAACTATTAACCTTGCATGCTT 480
DB 421 GTGAGAGGATCATCCAGATGCACTCATTCAGACAGAACTATTAACCTTGCATGCTT 480
QY 481 GGTGAGACAGAAATATACATGTTTATTTCTGAAATCCCAAACTATCATAGAGCA 540
DB

DB 481 GGTGAGACAGAAATATACATGTTTATTTCTGAAATCCCAAACTATCATAGAGCA 540
QY 541 GCAGTCTTAATGCTCTCTCGAAGCCCTCTTTTGGATCTTTTTCAGGCAACACTGAGCTAT 600
DB 541 GCAGTCTTAATGCTCTCTCGAAGCCCTCTTTTGGATCTTTTTCAGGCAACACTGAGCTAT 600
QY 601 GGAATGTTATCTCGAGAAAGAACTATTAAGAGAAAGAAACCGATTTGGAACTCGGA 660
DB 601 GGAATGTTATCTCGAGAAAGAACTATTAAGAGAAAGAAACCGATTTGGAACTCGGA 660
QY 661 ATTGCTTTTACAGTATATCAACCAAGGAAGTGGAACTATTTCTGTTTCAAGCCGGTAGTGA 720
DB 661 ATTGCTTTTACAGTATATCAACCAAGGAAGTGGAACTATTTCTGTTTCAAGCCGGTAGTGA 720
QY 721 ATTTATCACGTAAAAGATGGAGGCCACCAAGGATTTTACCAACAACTTTTAAGGCCCAAT 780
DB 721 ATTTATCACGTAAAAGATGGAGGCCACCAAGGATTTTACCAACAACTTTTAAGGCCCAAT 780
QY 781 CTAGTGGAAACTAGTTTGTCTCCCAACATACCGATGGATCCAAATTTATGCCCGCTGATCCA 840
DB 781 CTAGTGGAAACTAGTTTGTCTCCCAACATACCGATGGATCCAAATTTATGCCCGCTGATCCA 840
QY 841 GACTGGATTTGCTTTTATACATAGCAACGATATTTTGGATATCTAATCATCTAACCAAGAA 900
DB 841 GACTGGATTTGCTTTTATACATAGCAACGATATTTTGGATATCTAATCATCTAACCAAGAA 900
QY 901 GAAAGGAGACTCACTTTATGCAATGAGCTAGCCACATGCAATGCAATGCAATGCAATGCAATG 960
DB 901 GAAAGGAGACTCACTTTATGCAATGAGCTAGCCACATGCAATGCAATGCAATGCAATGCAATG 960
QY 961 GCTGGAGTCCGCTACCTTTGTTCTCCAAAGAAATTTGATAGATATTTCTGGCTATTTGGTGG 1020
DB 961 GCTGGAGTCCGCTACCTTTGTTCTCCAAAGAAATTTGATAGATATTTCTGGCTATTTGGTGG 1020
QY 1021 TGTCCAAAGCTGAAACCAACTCCAGTGGTGTGTAATAATTTCTAGAAATTTCTATATGAAGAA 1080
DB 1021 TGTCCAAAGCTGAAACCAACTCCAGTGGTGTGTAATAATTTCTAGAAATTTCTATATGAAGAA 1080
QY 1081 AATGATGAATCTGAGTGGAAATTTATTCATGTTTACATCCCTCTATGTTGGAAACCAAGGAGG 1140
DB 1081 AATGATGAATCTGAGTGGAAATTTATTCATGTTTACATCCCTCTATGTTGGAAACCAAGGAGG 1140
QY 1141 GCAGATTCATTCGGTTATCTCTAAACAGGTACAGCAAAATCCTTAAAGTCACCTTTTAAAGATG 1200
DB 1141 GCAGATTCATTCGGTTATCTCTAAACAGGTACAGCAAAATCCTTAAAGTCACCTTTTAAAGATG 1200
QY 1201 TCAGAAATATGATTTGATGCTGAAAGAAAGATCATAGATGTCTAGATGATGATGATGATGATG 1260
DB 1201 TCAGAAATATGATTTGATGCTGAAAGAAAGATCATAGATGTCTAGATGATGATGATGATGATG 1260
QY 1261 CAACCTTTTGGAGATTTCTATTGAAAGAGTTGAAATATTTGCGCAGAGCTGGATGACTCCT 1320
DB 1261 CAACCTTTTGGAGATTTCTATTGAAAGAGTTGAAATATTTGCGCAGAGCTGGATGACTCCT 1320
QY 1321 GAGGAAATATGCTTTGGTCCATCTCTACTAGATCGTCCAGACTCGCTTACAGATAGTG 1380
DB 1321 GAGGAAATATGCTTTGGTCCATCTCTACTAGATCGTCCAGACTCGCTTACAGATAGTG 1380
QY 1381 TTGATCTCACCTGAATTTATTTTATCCAGTGAAGATGATGATGATGATGATGATGATGATGATG 1440
DB 1381 TTGATCTCACCTGAATTTATTTTATCCAGTGAAGATGATGATGATGATGATGATGATGATGATG 1440
QY 1441 ATTGAGTCAGTGGCTGCTGCTGAGCGCATTAATTTCTATGATGATGATGATGATGATGATGATG 1500
DB 1441 ATTGAGTCAGTGGCTGCTGCTGAGCGCATTAATTTCTATGATGATGATGATGATGATGATGATG 1500
QY 1501 TGGATTAATTCATGACATCTTTTCTGAGCGCATTAATTTCTGATGATGATGATGATGATGATGATG 1560
DB 1501 TGGATTAATTCATGACATCTTTTCTGAGCGCATTAATTTCTGATGATGATGATGATGATGATGATG 1560
QY 1561 TTTATTTTGGCTCTGAATGCAAAAACAGGTTTCCGTCATTTTATATCAAAATTTACATCTATT 1620
DB 1561 TTTATTTTGGCTCTGAATGCAAAAACAGGTTTCCGTCATTTTATATCAAAATTTACATCTATT 1620

Mon May 8 13:41:52 2006

Qy	1621	TTAAGGAAAGCAATATAAAGCATCCAGTGGTGGCTGCTCCTCAAGTGATTTCAAG	1680
Db	1621	TTAAGGAAAGCAATATAAAGCATCCAGTGGTGGCTGCTCCTCAAGTGATTTCAAG	1680
Qy	1681	TGTCCTATCAAGAGGAGATAGCAATTACCAAGTGGTGAATGGGAAGTTCTTGGCCGGCAT	1740
Db	1681	TGTCCTATCAAGAGGAGATAGCAATTACCAAGTGGTGAATGGGAAGTTCTTGGCCGGCAT	1740
Qy	1741	GGATCTAATATCAAGTTGATGAAGTCAGAAGGCTGGTATATTTGAAGGCCACCAAGAC	1800
Db	1741	GGATCTAATATCAAGTTGATGAAGTCAGAAGGCTGGTATATTTGAAGGCCACCAAGAC	1800
Qy	1801	TCCCCCTTAGAGCATCACCTGTACGTAGTCAAGTTACGTAAATCCTGGAGAGGTGACAAGG	1860
Db	1801	TCCCCCTTAGAGCATCACCTGTACGTAGTCAAGTTACGTAAATCCTGGAGAGGTGACAAGG	1860
Qy	1861	CTGACTGACCGTGGCTACTCACATTTCTTGTGCTGCATCAGTCAGCACTGTGACTTCTTTTATA	1920
Db	1861	CTGACTGACCGTGGCTACTCACATTTCTTGTGCTGCATCAGTCAGCACTGTGACTTCTTTTATA	1920
Qy	1921	AGTAAGTATAGTAACAGAGAATCCACACTGTGTGCTCCCTTTACAAGCTATCAAGTCTT	1980
Db	1921	AGTAAGTATAGTAACAGAGAATCCACACTGTGTGCTCCCTTTACAAGCTATCAAGTCTT	1980
Qy	1981	GAAGATGACCCCACTTGCAGAACAAAGGAATTTTGGGCCACCATTTTGGATTTCAGCAGGT	2040
Db	1981	GAAGATGACCCCACTTGCAGAACAAAGGAATTTTGGGCCACCATTTTGGATTTCAGCAGGT	2040
Qy	2041	CCTCTTCTGACTATATCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGATTTACA	2100
Db	2041	CCTCTTCTGACTATATCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGATTTACA	2100
Qy	2101	TTGTATGGGATGCTCTACAGCCTCATGATCTACAGCCTGGAAGAAATATCCTACTGTG	2160
Db	2101	TTGTATGGGATGCTCTACAGCCTCATGATCTACAGCCTGGAAGAAATATCCTACTGTG	2160
Qy	2161	CTGTTTCATATATGTGTGTC	2179
Db	2161	CTGTTTCATATATGTGTGTC	2179

Search completed: May 4, 2006, 05:19:47
Job time : 1778 secs

THIS PAGE BLANK (USPTO)


GenCore version 5.1.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model
Run on: May 4, 2006, 05:20:06 ; Search time 11795 Seconds
(without alignments)
123776.066 Million cell updates/sec

Title: US-10-825-632-2
Perfect score: 3120
Sequence: 1 aagtgttaagctccgagg.....agaattactaaaaaaaaa 3120

Scoring table:  OMEGA-NUC
Gapop 60.0 , Gapext 60.0

Searched: 41078325 seqs, 23393541228 residues

Word size: 
Total number of hits satisfying chosen parameters: 82156240

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : EST:*

- 1: gb_est1.*
- 2: gb_est2.*
- 3: gb_est3.*
- 4: gb_est4.*
- 5: gb_est5.*
- 6: gb_est6.*
- 7: gb_est7.*
- 8: gb_est8.*
- 9: gb_est9.*
- 10: gb_gss2.*
- 11: gb_gss3.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1664	53.3	2292	4	CR609512
2	966	31.0	2649	10	AY411615
3	782	25.1	1292	4	AF175225
4	769	24.6	1042	3	BM557438
5	755	24.2	884	7	CR985650
6	726	23.3	910	2	BG479035
7	725	23.2	985	3	BQ068650
8	724	23.2	968	5	BQ671635
9	723	23.2	910	5	BQ675260
10	722	23.1	1265	5	AF176779
11	708	22.7	735	1	AL043338
12	700	22.4	753	3	AL040398
13	690	22.1	835	7	CR998849
14	687	22.0	742	7	CN427229
15	686	22.0	914	5	BX390898
16	685	22.0	724	7	CK000768
17	682	21.9	820	7	CR994557
18	676	21.7	864	1	AL542617
19	674	21.6	2649	10	AY411616
20	672	21.5	1002	2	BE733691
21	671	21.5	671	8	CX787143
22	663	21.2	804	5	BU687687

23	660	21.2	993	2	BG259714
24	652	20.9	910	5	BX372276
25	647	20.7	789	2	BG709118
26	644	20.6	669	2	BG390103
27	641	20.5	881	7	CR990494
28	639	20.5	708	5	BU622228
29	637	20.4	722	5	BU631054
30	627	20.1	853	2	BI223892
31	626	20.1	815	6	CB990233
32	626	20.1	826	3	BI553405
33	623	20.0	695	2	BG715866
34	614	19.7	634	1	AI819365
35	607	19.5	957	5	BQ675006
36	598	19.2	612	7	CV024163
37	596	19.1	596	6	CB153590
38	581	18.6	632	5	BX645431
39	580	18.6	1041	2	BI084090
40	576	18.5	627	2	BG752328
41	562	18.0	760	1	AI917735
42	561	18.0	678	7	CN427235
43	558	17.9	673	1	AW303607
44	555	17.8	788	2	BG163397
45	548	17.6	553	7	CN427230

ALIGNMENTS

RESULT 1
CR609512
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REMARK
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
ORIGIN

2292 bp mRNA linear HTC 21-JUL-2004
full-length cDNA clone CS0DL005YD02 of B cells (Ramos cell line)
Cot 25-normalized of Homo sapiens (human).
CR609512
CR609512.1 GI:50490319
HTC; CNSLT_CDNA.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 2292)
Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
2 (bases 1 to 2292)
Genoscope.
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
Location/Qualifiers
1. .2292
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DL005YD02"
/tissue_type="B cells (Ramos cell line) Cot 25-normalized"
/plasmid="pCMVSPORT_6"

Query Match 53.3%; Score 1664; DB 4; Length 2292;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1714; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 472 GCCATGCTGCTGAGACACAGAGAAATACACTGTTTATTCTGAAATTCCTCAAACTATC 531
Db 11 GCCATGCTGCTGAGACACAGAGAAATACACTGTTTATTCTGAAATTCCTCAAACTATC 70
QY 532 AATAGACGACAGCTTAATGCTCTCTCGAAGCCCTCTTTGGATCTTTTTCAGGGAACA 591
Db 71 AATAGACGACAGCTTAATGCTCTCTCGAAGCCCTCTTTGGATCTTTTTCAGGGAACA 130
QY 592 CTGCACTATGGAATGTAATCTCGAGAAGAAGAACTATTAAAGAGAAGAAAACGCAATTGGA 651
Db 131 CTGCACTATGGAATGTAATCTCGAGAAGAAGAACTATTAAAGAGAAGAAAACGCAATTGGA 190
QY 652 ACAGTCGGAATGCTCTTACGATTAATCAACGAAGAGTGGAACTTTCTGTTTCAAGCC 711
Db 191 ACAGTCGGAATGCTCTTACGATTAATCAACGAAGAGTGGAACTTTCTGTTTCAAGCC 250
QY 712 GGTAGTGGAAATTTATCACTGTAAGAGTGGAGGCCACAGGATTTTACGCAACAACTTTTA 771
Db 251 GGTAGTGGAAATTTATCACTGTAAGAGTGGAGGCCACAGGATTTTACGCAACAACTTTTA 310
QY 772 AGGCCCAATCTAGTGGAAATCTAGTGTCTCCAAACATACGATGGATCCAAAATTTATGCC 831
Db 311 AGGCCCAATCTAGTGGAAATCTAGTGTCTCCAAACATACGATGGATCCAAAATTTATGCC 370
QY 832 GCTGATCCAGACTGGATTTGCTTTTATACATAGCAACGATATTGATGATATCTAAACATGTA 891
Db 371 GCTGATCCAGACTGGATTTGCTTTTATACATAGCAACGATATTGATGATATCTAAACATGTA 430
QY 892 ACCAGAGAAGAAAGGAGACTCACTTATGTGCAACATGAGCTAGCCACATGGAGAAGAT 951
Db 431 ACCAGAGAAGAAAGGAGACTCACTTATGTGCAACATGAGCTAGCCACATGGAGAAGAT 490
QY 952 GCCAGATCAGCTGGAGTGGTCTGCTTCTCCGAAGAAATTTGATAGATTTCTGGC 1011
Db 491 GCCAGATCAGCTGGAGTGGTCTGCTTCTCCGAAGAAATTTGATAGATTTCTGGC 550
QY 1012 TATTGGTGGTCTCAAAAGCTGAACAACTCCAGTGGTGTGTAATTTCTTAGAATTTCTA 1071
Db 551 TATTGGTGGTCTCAAAAGCTGAACAACTCCAGTGGTGTGTAATTTCTTAGAATTTCTA 610
QY 1072 TATGAAGAAATGATGAATCTGAGTGGAAATTTATCATGTTACATCCCTCTATGTTGAA 1131
Db 611 TATGAAGAAATGATGAATCTGAGTGGAAATTTATCATGTTACATCCCTCTATGTTGAA 670
QY 1132 ACAAGGAGGACAGATTCATTCCGTTATCTCTAAACAGGTACAGCAAAATCTCTAAAGTCAC 1191
Db 671 ACAAGGAGGACAGATTCATTCCGTTATCTCTAAACAGGTACAGCAAAATCTCTAAAGTCAC 730
QY 1192 TTTAAGATGTCAGAAATAATGATGATGCTGAAGGAAGGATCATAGATGTCATAGATAAG 1251
Db 731 TTTAAGATGTCAGAAATAATGATGATGCTGAAGGAAGGATCATAGATGTCATAGATAAG 790
QY 1252 GAACATAATCAACCTTTTGAGATTTCTATTGGAAGGAGTTGAATATATTGCGAGAGCTGGA 1311
Db 791 GAACATAATCAACCTTTTGAGATTTCTATTGGAAGGAGTTGAATATATTGCGAGAGCTGGA 850
QY 1312 TGGACTCTCAGGGAATAATGCTTGGTCCATCTCTACTAGATCGCTCCAGACTCGCCTA 1371
Db 851 TGGACTCTCAGGGAATAATGCTTGGTCCATCTCTACTAGATCGCTCCAGACTCGCCTA 910
QY 1372 CAGATAGTGTGATCTCAGCTGAATTTATTCAGGATAGAGATGATGTTATGGAAG 1431
Db 911 CAGATAGTGTGATCTCAGCTGAATTTATTCAGGATAGAGATGATGTTATGGAAG 970
QY 1432 CAGAGACTCATAGTCACTGCTGATTTCTGTGAGCCCACTAATATCTATGGAAGAAACA 1491
Db 971 CAGAGACTCATAGTCACTGCTGATTTCTGTGAGCCCACTAATATCTATGGAAGAAACA 1030
QY 1492 ACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTTTTCCCAAGATCAGGAAG 1551
Db 1031 ACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTTTTCCCAAGATCAGGAAG 1090
QY 1552 GAAATTGAGTTATTTTCCCTCTCTGAATGCAAAACAGGTTTCCGTCATTTTATACAAAAT 1611
```

```
Db 1091 GAAATTGAGTTATTTTCCCTCTGAATGCAAAACAGGTTTCCGTCATTTATACAAAAT 1150
QY 1612 ACATCTATTTTAAAGGAAGCAAAATATAACCGATCCAGTGGTGGGCTGCCCTGCCAAGT 1671
Db 1151 ACATCTATTTTAAAGGAAGCAAAATATAACCGATCCAGTGGTGGGCTGCCCTGCCAAGT 1210
QY 1672 GATTTCAGAGTGTCTCTATCAAGAGGAGATAGCAATTTACAGTGGTGAATGGGAAGTTCTT 1731
Db 1211 GATTTCAGAGTGTCTCTATCAAGAGGAGATAGCAATTTACAGTGGTGAATGGGAAGTTCTT 1270
QY 1732 GCGCGGATCGATCTAATATCAAGTGTGATGAAGTCAGAAGGCTGGTATATTTTGAAGGC 1791
Db 1271 GCGCGGATCGATCTAATATCAAGTGTGATGAAGTCAGAAGGCTGGTATATTTTGAAGGC 1330
QY 1792 ACCAAAGACTCCCTTTTAGAGCATCACCTGTACGTAGTCAGTTACGTAAATCTCTGGAGAG 1851
Db 1331 ACCAAAGACTCCCTTTTAGAGCATCACCTGTACGTAGTCAGTTACGTAAATCTCTGGAGAG 1390
QY 1852 GTGACAAAGGCTGACTGACCGTGGCTACTCACATTTCTGTCATCAGTCAGCACTGTGAC 1911
Db 1391 GTGACAAAGGCTGACTGACCGTGGCTACTCACATTTCTGTCATCAGTCAGCACTGTGAC 1450
QY 1912 TTCTTTATAGTAAAGTATAGTAACCAAGAGATCCACACTGTGTGCTTTCACAGCTA 1971
Db 1451 TTCTTTATAGTAAAGTATAGTAACCAAGAGATCCACACTGTGTGCTTTCACAGCTA 1510
QY 1972 TCAAGTCTCGAAGTACGCCAACTTGCACAAACAAAGAAATTTTGGGCACTTTTGGAT 2031
Db 1511 TCAAGTCTCGAAGTACGCCAACTTGCACAAACAAAGAAATTTTGGGCACTTTTGGAT 1570
QY 2032 TCAGCAGGCTCTCTCTGACTATACTCTCCAGAAATTTTCTCTTTTGAAGTACTACT 2091
Db 1571 TCAGCAGGCTCTCTCTGACTATACTCTCCAGAAATTTTCTCTTTTGAAGTACTACT 1630
QY 2092 GGATTTACATGTPATGGAGTCTCTACAGCTCAATGATCTACAGCTCGGAAAGAAATAT 2151
Db 1631 GGATTTACATGTPATGGAGTCTCTACAGCTCAATGATCTACAGCTCGGAAAGAAATAT 1690
QY 2152 CCTACTGTGCTGTTCATATATATGCTGCTCCTCAGGT 2186
Db 1691 CCTACTGTGCTGTTCATATATATGCTGCTCCTCAGGT 1725

RESULT 2
AY411615 2649 bp DNA linear GSS 16-DEC-2003
LOCUS Homo sapiens DPP8 gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION Genomic survey sequence.
ACCESSION AY411615
VERSION AY411615.1 GI:39767583
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo (base 1 to 2649)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferrieres,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
Science 302 (5652), 1960-1963 (2003)
JOURNAL
PUBMED 14671302
REFERENCE 2. (bases 1 to 2649)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferrieres,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Direct Submission
TITLE Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
```

Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering them based on alignment.

FEATURES
source Location/Qualifiers
1..2649
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..>2649
/gene="DPP8"
/locus_tag="HCM4283"

ORIGIN
Query Match 31.0%; Score 966; DB 10; Length 2649;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1016; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 214 ATGGCAGCAGCAATGAAACAGACAGACTGGGTGTGAGATATTTGAAACTGCGACTGT 273
DB 1 ATGGCAGCAGCAATGAAACAGACAGACTGGGTGTGAGATATTTGAAACTGCGACTGT 60

QY 274 GAGGAGAAATTAATACAGATCGGCTTAAATTGAGCGCTTTTATGTTGACGGTAT 333
DB 61 GAGGAGAAATTAATACAGATCGGCTTAAATTGAGCGCTTTTATGTTGACGGTAT 120

QY 334 TCCTGAGTCAGCTTAAAGCTGCTGCCGATACCCAGAAATATCATGCTACATGATG 393
DB 121 TCCTGAGTCAGCTTAAAGCTGCTGCCGATACCCAGAAATATCATGCTACATGATG 180

QY 394 GCTAAGCCACCATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGACCTCATCA 453
DB 181 GCTAAGCCACCATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGACCTCATCA 240

QY 454 GACAGAATCTTATACCTGCCATGCTGGTGAGAACAGAGAAATACACTGTTTATTCT 513
DB 241 GACAGAATCTTATACCTGCCATGCTGGTGAGAACAGAGAAATACACTGTTTATTCT 300

QY 514 GAAATCCCAAACTATCAATAGACGACGCTTAAATGCTCTCTCGAAGCCTCTTTTG 573
DB 301 GAAATCCCAAACTATCAATAGACGACGCTTAAATGCTCTCTCGAAGCCTCTTTTG 360

QY 574 GATCTTTTTCAGGCAACACTGGAATGGAATGATTCGAGAGGAAGAACTATTAAAGA 633
DB 361 GATCTTTTTCAGGCAACACTGGAATGGAATGATTCGAGAGGAAGAACTATTAAAGA 420

QY 634 GAAAGAAAACCGCAATGGAACAGTCGGAATGCTTACAGATTATCACCAAGGAAGTGA 693
DB 421 GAAAGAAAACCGCAATGGAACAGTCGGAATGCTTACAGATTATCACCAAGGAAGTGA 480

QY 694 AATTTCTGTTCAAGCCGGTAGTGGAATTTATCAGTAAAGATGGAGGCCACAAAGGA 753
DB 481 AATTTCTGTTCAAGCCGGTAGTGGAATTTATCAGTAAAGATGGAGGCCACAAAGGA 540

QY 754 TTTACGCAACAACTTTAAGGCCCAATCTAGTGGAACCTAGTTGTCCCAACATACGGATG 813
DB 541 TTTACGCAACAACTTTAAGGCCCAATCTAGTGGAACCTAGTTGTCCCAACATACGGATG 600

QY 814 GATCCAAAATTTATGCCCGCTGATCCAGACTGGATGCTTTTATACATAGCAAGATTT 873
DB 601 GATCCAAAATTTATGCCCGCTGATCCAGACTGGATGCTTTTATACATAGCAAGATTT 660

QY 874 TGGATNTCTAACATCGTAACACAGAGAAAGGAGACTCACTTATGTGCACAATGAGCTA 933
DB 661 TGGATNTCTAACATCGTAACACAGAGAAAGGAGACTCACTTATGTGCACAATGAGCTA 720

QY 934 GCCAACATGGAAGAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTTCTCCAAAGAGAA 993
DB 721 GCCAACATGGAAGAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTTCTCCAAAGAGAA 780

QY 994 TTTGATAGATATTTCTGGCTATTTGGTGTGTCAAAAGCTGAAAACAACTCCCAAGTGGT 1053
DB 781 TTTGATAGATATTTCTGGCTATTTGGTGTGTCAAAAGCTGAAAACAACTCCCAAGTGGT 840

QY 1054 AAAATCTTCTAGATTTCTATATGAAAGAAATGATGATCTCAGGTGAAATATTATCATGTT 1113
DB 841 AAAATCTTCTAGATTTCTATATGAAAGAAATGATGATCTCAGGTGAAATATTATCATGTT 900

QY 1114 ACATCCCTATGTTGGAACCAAGGAGGCGAGATTCAATTCCTGTTATCTCTAAACAGGTACA 1173
DB 901 ACATCCCTATGTTGGAACCAAGGAGGCGAGATTCAATTCCTGTTATCTCTAAACAGGTACA 960

QY 1174 GCAAATCCTAAAGTCACATTTTAAAGATGTCAGAAATAATGATGCTGAAGGAAG 1230
DB 961 GCAAATCCTAAAGTCACATTTTAAAGATGTCAGAAATAATGATGCTGAAGGAAG 1017

RESULT 3
AF175225 1292 bp mRNA linear HTC 01-AUG-2003
LOCUS Homo sapiens tissue-type aorta MSTP135 mRNA, complete cds.
DEFINITION AF175225
ACCESSION AF175225
VERSION AF175225.1 GI:33338055
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 1292)
AUTHORS Zhao, B., Xu, H.S., Tong, Y.K., Sheng, H., Qin, B.M., Liu, Y.Q., Liu, B., Wang, X.Y., Zhang, Q., Song, L., Gao, Y., Zhang, C.L., Ye, J., Ji, X.J., Liu, B.H., Lu, H., Chen, J.Z., Cai, M.Q., Zheng, W.Y., Teng, C.Y., Liu, Q., Yu, J.T., Lin, J., Gong, Q., Zhang, A.M., Gao, R.L. and Hui, R.T.
Direct Submission
Submitted (04-AUG-1999) Molecular Medicine Center for Cardiovascular Disease, Cardiovascular Institute, CNMS & PUMC, 167, Bei Li Shi Lu, Beijing 100037, P.R. China
JOURNAL Location/Qualifiers
FEATURES
source 1..1292
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="aorta"
142..1023
/codon_start=1
/product="MSTP135"
/protein_id="AAQ13650.1"
/db_xref="GI:33338056"
/translation="MTQLAKORNFPPFWIQVLFLLTLQKFSLLKVLDDLHMCQCS
TSPHSGTGLERNILLXLFYGPQKVLVNNRFGKVFRLNLPFLGCVVXGIDNRGS
CPKGLFEGAFKFMQIEIDQVEGLQYLSRVDFIDLRVGIHGWYSYGLSLMAL
MQRSDIFRVAIAGPVTLWIFDTGTYTERYMGHPDQNEQGYILGSVAMQAEKPSPE
RLLLHGFIDENVHFAHTSILLSFLVRAGKPYDILQIYPOERKHSIRVPESGEHVELHL
HYLQENLGSRIAALKVI"
ORIGIN
Query Match 25.1%; Score 782; DB 4; Length 1292;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 782; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2325 TAAATGGTCAATAGAAATTCAGCATCAGGTGGAAGACTCCAAATATCTAGCTTCTCG 2384
DB 486 TAAATGGTCAATAGAAATTCAGCATCAGGTGGAAGACTCCAAATATCTAGCTTCTCG 545

QY 2385 ATATGATTTCAATGACTTAGATCGTGTGCGCATCCACGGCTGCTTATGAGGATACCT 2444
DB 546 ATATGATTTCAATGACTTAGATCGTGTGCGCATCCACGGCTGCTTATGAGGATACCT 605

QY 2445 CTCCTCATGGCATTATGACAGGTCAGATATCTTCAGGTGCTATTGCTGGGGCCCC 2504
DB 606 CTCCTCATGGCATTATGACAGGTCAGATATCTTCAGGTGCTATTGCTGGGGCCCC 665

QY 2505 AGTCACCTCTGTGGATCTTCTATGATACAGATACACGGAACGTTATATGGGTACACCTGA 2564
DB 666 AGTCACCTCTGTGGATCTTCTATGATACAGATACACGGAACGTTATATGGGTACACCTGA 725

```
QY 2565 CCAGAAATGAACAGCGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCCTC 2624
Db 726 CCAGAAATGAACAGCGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCCTC 785
QY 2625 TGAACCAAAATCGTTTACTGCTCTTACATGGTTTCTGGATGAGATGTCATTTTGACA 2684
Db 786 TGAACCAAAATCGTTTACTGCTCTTACATGGTTTCTGGATGAGATGTCATTTTGACA 845
QY 2685 TACCAGTATATTACTGATGTTTATGATGAGGCTGGAAGCCATATGATTTACAGATCTA 2744
Db 846 TACCAGTATATTACTGATGTTTATGATGAGGCTGGAAGCCATATGATTTACAGATCTA 905
QY 2745 TCCTCAGAGAGACACAGCATAGAGTTCTGATCGGGAGAACATTAAGTAACTGCATCT 2804
Db 906 TCCTCAGAGAGACACAGCATAGAGTTCTGATCGGGAGAACATTAAGTAACTGCATCT 965
QY 2805 TTTCGACTACTCTCAAGAAAACCTTGGATCAGGTATTCCTCTTAAAGTGATATAAT 2864
Db 966 TTTCGACTACTCTCAAGAAAACCTTGGATCAGGTATTCCTCTTAAAGTGATATAAT 1025
QY 2865 TTGACCTGTGTAGAACTCTCTGGTATACACTGGCTATTTAAACCAAAATGAGAGTTTAA 2924
Db 1026 TTGACCTGTGTAGAACTCTCTGGTATACACTGGCTATTTAAACCAAAATGAGAGTTTAA 1085
QY 2925 CAACAGAAAACACAGAAATGATCATCACATTTTGATACCTGCCATGTAAACATCTACTCCT 2984
Db 1086 CAACAGAAAACACAGAAATGATCATCACATTTTGATACCTGCCATGTAAACATCTACTCCT 1145
QY 2985 GAAATAAATGTGTGCCATGAGGGGTCTACGGTTTGTGTAGTAACTAACTACCTTAA 3044
Db 1146 GAAATAAATGTGTGCCATGAGGGGTCTACGGTTTGTGTAGTAACTAACTACCTTAA 1205
QY 3045 CCCACATGCTCAAAATCAAAATGATACATATTTCTGAGAGACCCGCAATACCATTAAGAA 3104
Db 1206 CCCACATGCTCAAAATCAAAATGATACATATTTCTGAGAGACCCGCAATACCATTAAGAA 1265
QY 3105 TT 3106
Db 1266 TT 1267
```

```
RESULT 4
LOCUS BM557438 1042 bp mRNA linear EST 20-FEB-2002
DEFINITION AGENCOURT_5578992 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:5466759
S', mRNA sequence.
ACCESSION BM557438
VERSION 1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 1042)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue procurement: DCTD/BRP
cDNA library preparation: Rubin Laboratory
cDNA library arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
plate: L1C11967 row: k column: 16
High quality sequence stop: 697.
1..1042
Location/Qualifiers
/molecule="mRNA"
/molecule="Homo sapiens"
source
```

```
/db_xref="taxon:9606"
/clone="IMAGE:5466759"
/tissue_type="amelanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 41"
/notes="Organ: skin; Vector: pOTB7; Site: 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."
```

Query Match	Best Local Similarity	Score	DB 3	Length	1042
Matches	819	Conservative	99.9%	Pred. No.	0
				Mismatches	1
				Indels	0
				Gaps	0
QY	1723	GAAGTTCTTGGCGGCATGATCTAATATCCAGTTGATGAGTCCAGAGGCTGGTATAT	1782		
Db	1	GAAGTTCTTGGCGGCATGATCTAATATCCAGTTGATGAGTCCAGAGGCTGGTATAT	60		
QY	1783	TTTGAAGGACCAAGACCTCCCTTTAGAGCATACCTGTAGTACGTTACGTTACGTTAAAT	1842		
Db	61	TTTGAAGGACCAAGACCTCCCTTTAGAGCATACCTGTAGTACGTTACGTTACGTTAAAT	120		
QY	1843	CCTGGAGAGTGACAAAGGCTGACTGACCGTGGCTACTCACATTTCTGCTGATCAGTCAG	1902		
Db	121	CCTGGAGAGTGACAAAGGCTGACTGACCGTGGCTACTCACATTTCTGCTGATCAGTCAG	180		
QY	1903	CATGTGACTCTCTTATAGTAAGTATAGTAACAGAGAAATCCACACTGTGTGTCCTTT	1962		
Db	181	CATGTGACTCTCTTATAGTAAGTATAGTAACAGAGAAATCCACACTGTGTGTCCTTT	240		
QY	1963	TACAAGCTATCAAGTCTCTGAGATGACCCAACTTTGCAAAAACAAAGGAAATTTGGGCCACC	2022		
Db	241	TACAAGCTATCAAGTCTCTGAGATGACCCAACTTTGCAAAAACAAAGGAAATTTGGGCCACC	300		
QY	2023	ATTTTGGATTACAGCAGGTCCTCTTCTGACTATATCTCTCCAGAAATTTCTCTTTTGA	2082		
Db	301	ATTTTGGATTACAGCAGGTCCTCTTCTGACTATATCTCTCCAGAAATTTCTCTTTTGA	360		
QY	2083	AGTACTACTGGAATTTTACATTTGATGGGATGCTCTACAGGCTCATGATCTACAGCCTGGA	2142		
Db	361	AGTACTACTGGAATTTTACATTTGATGGGATGCTCTACAGGCTCATGATCTACAGCCTGGA	420		
QY	2143	AAGAAATATCTTACTGCTGCTGTTTATATATGTTGGTCTCTCAGTGCAATTTGGTAAAT	2202		
Db	421	AAGAAATATCTTACTGCTGCTGTTTATATATGTTGGTCTCTCAGTGCAATTTGGTAAAT	480		
QY	2203	CGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTAGCCTCTCTAGTTATGTTGTT	2262		
Db	481	CGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTAGCCTCTCTAGTTATGTTGTT	540		
QY	2263	GTAGTGATAGACACAGGGGATCTCTGTCACCGAGGCTTAAATTTGAAGCGCCTTTTAA	2322		
Db	541	GTAGTGATAGACACAGGGGATCTCTGTCACCGAGGCTTAAATTTGAAGCGCCTTTTAA	600		
QY	2323	TATAAAATGGGTCAAAATAGAAAATTTGACATCAGGTGGAAGGACTCCAATATCTAGTTCT	2382		
Db	601	TATAAAATGGGTCAAAATAGAAAATTTGACATCAGGTGGAAGGACTCCAATATCTAGTTCT	660		
QY	2383	CGATATGATTTTCAITGACTTAGATCGTGTGGGATCCACGGCTGGTCTTATGGAGATAC	2442		
Db	661	CGATATGATTTTCAITGACTTAGATCGTGTGGGATCCACGGCTGGTCTTATGGAGATAC	720		
QY	2443	CTCTCCCTGATGGCATTAAATGACAGAGTCAAGATATCTTACGGGTTGCTATTGCTGGGCC	2502		
Db	721	CTCTCCCTGATGGCATTAAATGACAGAGTCAAGATATCTTACGGGTTGCTATTGCTGGGCC	780		
QY	2503	CCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGG	2542		


```
Plate: LLCMI430 row: e column: 03
High quality sequence stop: 837.
FEATURES
  source
    1..910
      /organism="Homo sapiens"
      /mol_type="mrna"
      /db_xref="taxon:9606"
      /clone="IMAGE:4649834"
      /tissue_type="choriocarcinoma"
      /lab_host="DH10B (phage-resistant)"
      /clone_lib="NIH_MGC_21"
      /note="Organ: placenta; Vector: pOTB7; Site: 1: XhoI;
      Site_2: EcoRI; cDNA made by oligo-dt priming.
      Directionally cloned into EcoRI/XhoI sites using the
      following 5' adaptor: GGCACGAG(G). Size-selected >500bp
      for average insert size 1.8kb. Library constructed by
      Ling Hong in the laboratory of Gerald M. Rubin (University
      of California, Berkeley) using ZAP-cDNA synthesis kit
      (Stratagene) and Superscript II RT (Life Technologies)."
```

ORIGIN

Query Match 23.3%; Score 726; DB 2; Length 910;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 726; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2193 GGTGAATAATCGGTTTAAAGGAGTCAAGTATTTCGGCTTGAATACCCCTAGCCTCTCTAGG 2252
Db 1 GGTGAATAATCGGTTTAAAGGAGTCAAGTATTTCGGCTTGAATACCCCTAGCCTCTCTAGG 60

QY 2253 TTATGTGTTGTAGTGATAGACAACAGGGGATCCTGTCAACGAGGGCTTAAATTTGAAGG 2312
Db 61 TTATGTGTTGTAGTGATAGACAACAGGGGATCCTGTCAACGAGGGCTTAAATTTGAAGG 120

QY 2313 CGCCTTTAAATATAAATGGTCAATAGAAATTCAGCATCAGTGGAGGATCCCAATA 2372
Db 121 CGCCTTTAAATATAAATGGTCAATAGAAATTCAGCATCAGTGGAGGATCCCAATA 180

QY 2373 TCTAGCTTCGAGATAGATTTCATTCAGTCTAGATCGTGTGGCATCAACGGCTGTCTTA 2432
Db 181 TCTAGCTTCGAGATAGATTTCATTCAGTCTAGATCGTGTGGCATCAACGGCTGTCTTA 240

QY 2433 TGGAGGATACCTTCCTCGATGGCATTAATGCAGAGTTCAGATATCTTCAGGGTTCGTAT 2492
Db 241 TGGAGGATACCTTCCTCGATGGCATTAATGCAGAGTTCAGATATCTTCAGGGTTCGTAT 300

QY 2493 TGCTGGGGCCCGAGTCACCTCTGTGGATCTTCTATGATACAGGATACACGGAACGTTATAT 2552
Db 301 TGCTGGGGCCCGAGTCACCTCTGTGGATCTTCTATGATACAGGATACACGGAACGTTATAT 360

QY 2553 GGGTCAACCTGACCAAGATGAACAGGGCTATTACTTTAGGATCTGTGCCATGCAAGCAGA 2612
Db 361 GGGTCAACCTGACCAAGATGAACAGGGCTATTACTTTAGGATCTGTGCCATGCAAGCAGA 420

QY 2613 AAAAGTTCCTCTCTGAACCAAAATCGTTTACTGCTCTTACATGTTTCTTGGATGAGAAATGT 2672
Db 421 AAAAGTTCCTCTCTGAACCAAAATCGTTTACTGCTCTTACATGTTTCTTGGATGAGAAATGT 480

QY 2673 CCATTTTGACATACACAGTATATTACTCAGTTTTTTTGTAGGGCTGGAAGCCATATCA 2732
Db 481 CCATTTTGACATACACAGTATATTACTCAGTTTTTTTGTAGGGCTGGAAGCCATATCA 540

QY 2733 TTTACAGATCTATCTCTCAGGAGAGACAGCATAGAGTTCCTGATTCGGGAGCAATTA 2792
Db 541 TTTACAGATCTATCTCTCAGGAGAGACAGCATAGAGTTCCTGATTCGGGAGCAATTA 600

QY 2793 TGAACCTGCATCTTTTGCACTACCTTCAGRAAAACCTTGGATCATCGATTGCTCTTAA 2852
Db 601 TGAACCTGCATCTTTTGCACTACCTTCAGRAAAACCTTGGATCATCGATTGCTCTTAA 660

QY 2853 AGTGATATAATTTTGCACCTGTGAGAACTCTCTGGTATACATGGCTATTTTAAACCAATG 2912
Db 661 AGTGATATAATTTTGCACCTGTGAGAACTCTCTGGTATACATGGCTATTTTAAACCAATG 720

QY 2913 AGGAGG 2918
Db 721 AGGAGG 726

RESULT 7
BQ068650
LOCUS BQ068650.1 GI:19897696
DEFINITION 5', mRNA sequence.
ACCESSION BQ068650
VERSION BQ068650.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 985)
NIH-MGC http://mgc.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2035 row: 0 column: 08
High quality sequence stop: 677.
Location/Qualifiers
1..985
/organism="Homo sapiens"
/mol_type="mrna"
/db_xref="taxon:9606"
/clone="IMAGE:5802079"
/tissue_type="neuroblastoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_47"
/note="Organ: brain; Vector: pOTB7; Site 1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dt priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN

Query Match 23.2%; Score 725; DB 3; Length 985;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 725; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 63 TTCGCGCGCTGGTGTGTACCGGGCGCCGCCCGAGGAGCCACTGCAACGAGCCGGA 122
Db 1 TTCGCGCGCTGGTGTGTACCGGGCGCCGCCCGAGGAGCCACTGCAACGAGCCGGA 60

QY 123 GTGAGCG 182
Db 61 GTGAGCG 120

QY 183 CGGCT 242
Db 121 CGGCT 180

QY 243 GGGTGTGTGAGATATTGTGAAACTGCGGACTGTGAGGAGAAATATTGAAATCACAGGATCGGCC 302
Db 181 GGGTGTGTGAGATATTGTGAAACTGCGGACTGTGAGGAGAAATATTGAAATCACAGGATCGGCC 240

```
QY 303 TAAATGGAGCGCTTTTATGTTGAGCGGTATCTCTGGAGTCAGCTTAAAAAGCTGCTTGC 362
Db 241 TAAATGGAGCGCTTTTATGTTGAGCGGTATCTCTGGAGTCAGCTTAAAAAGCTGCTTGC 300
QY 363 CQATACGAGAAAATATCATGCTACATGATGGCTAGGTCAGGACCAACATGATTTTCATGTTGT 422
Db 301 CQATACGAGAAAATATCATGCTACATGATGGCTAGGTCAGGACCAACATGATTTTCATGTTGT 360
QY 423 GAAGAGGAATCATCCAGATGGACTCATTCAGACAGAACTATTACCTTGGCCATGCTCGG 482
Db 361 GAAGAGGAATCATCCAGATGGACTCATTCAGACAGAACTATTACCTTGGCCATGCTCGG 420
QY 483 TGAGAACAGAGAAAATACACTGTTTATCTGAAATTCCTCAAACTATCAATAGAGCAGC 542
Db 421 TGAGAACAGAGAAAATACACTGTTTATCTGAAATTCCTCAAACTATCAATAGAGCAGC 480
QY 543 AGTCTTAATGCTCTCTTGGAGCGCTCTTTTGGATCTTTTTCAGGCAACACTGGACTATGG 602
Db 481 AGTCTTAATGCTCTCTTGGAGCGCTCTTTTGGATCTTTTTCAGGCAACACTGGACTATGG 540
QY 603 AATGTATTCTCGAAGAAGAACTATTAAAGAGAAAGAAACGCATTGGAAACAGTCGGAAT 662
Db 541 AATGTATTCTCGAAGAAGAACTATTAAAGAGAAAGAAACGCATTGGAAACAGTCGGAAT 600
QY 663 TGCTTCTTACGATTATACCAAGAGAGTGGAACATTTCTGTTTCAAGCGCGTAGTGAAT 722
Db 601 TGCTTCTTACGATTATACCAAGAGAGTGGAACATTTCTGTTTCAAGCGCGTAGTGAAT 660
QY 723 TTATCAGTAAAGATGGAGGCGCACAGAGATTACGCAACACTTTAAAGGCCCAATCT 782
Db 661 TTATCAGTAAAGATGGAGGCGCACAGAGATTACGCAACACTTTAAAGGCCCAATCT 720
QY 783 AGTGG 787
Db 721 AGTGG 725
```

RESULT 8
B0671635
LOCUS
DEFINITION
AGENCOURT 8203887 NIH_MGC_102 Homo sapiens cDNA clone IMAGE:6255645
5', mRNA sequence.

ACCESSION
B0671635
VERSION
B0671635.1 GI:21782469
KEYWORDS
EST.
SOURCE
Homo sapiens (human)

ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE
1 (bases 1 to 968)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)

AUTHORS
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC

COMMENT
cDNA Library Arrayed by: Rubin Laboratory
cDNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>

Plate: L1CM2407 row: a column: 22
High quality sequence stop: 630.
Location/Qualifiers
1..968

FEATURES
source
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6255645"
/tissue_type="epidermoid carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"

/clone lib="NIH MGC 102"
/note="Organ: salivary gland; Vector: pOTB7; Site 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed
by Ling Hong in the laboratory of Gerald M. Rubin
(University of California, Berkeley) using ZAP-cDNA
synthesis kit (Stratagene) and Superscript II RT (Life
Technologies). Note: this is a NIH_MGC Library."

ORIGIN
Query Match 23.2%; Score 724; DB 5; Length 968;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 774; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1634 AATATAAACGATCCAGTGGTGGCTGCTCTCCAAAGTGATTTCAAGTGTCTCTCAAG 1693
Db 1 AATATAAACGATCCAGTGGTGGCTGCTCTCCAAAGTGATTTCAAGTGTCTCTCAAG 60
QY 1694 AGGAGATAGCAATTAACAGTGGTGAATGGGAAGTTCTTGGCCGGCATGGATCTAATATCC 1753
Db 61 AGGAGATAGCAATTAACAGTGGTGAATGGGAAGTTCTTGGCCGGCATGGATCTAATATCC 120
QY 1754 AAGTTGATAGTCAAGAGGCTGGTATATTTGAAGGACCAAGACTCCCTTTAGAGC 1813
Db 121 AAGTTGATAGTCAAGAGGCTGGTATATTTGAAGGACCAAGACTCCCTTTAGAGC 180
QY 1814 ATCACTGTAGTGTAGTGTAGTGTAAATCTCTGGAGAGGTGACAAAGGCTGACTGACCGTG 1873
Db 181 ATCACTGTAGTGTAGTGTAGTGTAAATCTCTGGAGAGGTGACAAAGGCTGACTGACCGTG 240
QY 1874 GCTACTCACATTTCTTGTGTCATCAGTGTGACCTTCTTTTATAAGTAAAGTATAGTA 1933
Db 241 GCTACTCACATTTCTTGTGTCATCAGTGTGACCTTCTTTTATAAGTAAAGTATAGTA 300
QY 1934 ACCAAGAAATCCACACTGTGTCTCTTTACAAGCTATCAAGTCTCTGAAGATGACCCAA 1993
Db 301 ACCAAGAAATCCACACTGTGTCTCTTTACAAGCTATCAAGTCTCTGAAGATGACCCAA 360
QY 1994 CTTGCAAAAACAAAGGAATTTTGGGCCACCATTTTGGATTTCAGCAGGTCCTCTCTCTGACT 2053
Db 361 CTTGCAAAAACAAAGGAATTTTGGGCCACCATTTTGGATTTCAGCAGGTCCTCTCTCTGACT 420
QY 2054 ATACTCTCTCAGAAATTTTCTCTTTTGAAGTACTACTGATTTACATTTATGGGATGC 2113
Db 421 ATACTCTCTCAGAAATTTTCTCTTTTGAAGTACTACTGATTTACATTTATGGGATGC 480
QY 2114 TCTCAAGCCTCATGATCTACAGCCTGGAAAGAAATATCTACTGTCTGTTCATATATG 2173
Db 481 TCTCAAGCCTCATGATCTACAGCCTGGAAAGAAATATCTACTGTCTGTTCATATATG 540
QY 2174 GTGGTCTCAGGTGTCAGTGTGGTGAATAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGA 2233
Db 541 GTGGTCTCAGGTGTCAGTGTGGTGAATAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGA 600
QY 2234 ATACCCTAGCCTCTCTAGGTTATGTGGTTGTAGTATAGACAAACAGGGGATCTGTGCACC 2293
Db 601 ATACCCTAGCCTCTCTAGGTTATGTGGTTGTAGTATAGACAAACAGGGGATCTGTGCACC 660
QY 2294 GAGGGCTTAAATTTGAAGGCGCTTTTAAATATAAATGGTCAAAATAGAAATTCAGCATC 2353
Db 661 GAGGGCTTAAATTTGAAGGCGCTTTTANATATAAATGGTCAAAATAGAAATTCAGCATC 720
QY 2354 AGGTGGAGGACTCCCAATATCTAGCTTCGATATGATTTTCACTTACCTAGATCG 2408
Db 721 AGGTGGAGGACTCCCAATATCTAGCTTCGATATGATTTTCACTTACCTAGATCG 775

RESULT 9
B0675260
LOCUS
DEFINITION
B0675260 910 bp mRNA linear EST 15-JUL-2002
AGENCOURT 8354972 NIH_MGC_102 Homo sapiens cDNA clone IMAGE:6275478
5', mRNA sequence.

ACCESSION BQ675260
VERSION BQ675260.1 GI:21786094
KEYWORDS EST.. sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 910)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: L1CM2458 row: 1 column: 07
High quality sequence stop: 618.
Location/Qualifiers
1..910
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clones="IMAGE:6275478"
/issue_type="epidermoid carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_102"
Site1: EcoRI; Salivary gland; Vector: pORF7; Site_1: XhoI;
Site2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed
by Ling Hong in the laboratory of Gerald M. Rubin
(University of California, Berkeley) using ZAP-cDNA
synthesis kit (Stratagene) and Superscript II RT (Life
Technologies). Note: this is a NIH_MGC Library."
ORIGIN
Query Match 23.2%; Score 723; DB 5; Length 910;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 823; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1635 ATATAACGATCCAGTGGTGGCTGCTCCTCAAGTGATTTCAAGTGCTCTATCAAGA 1694
DB 1 ATATAACGATCCAGTGGTGGCTGCTCCTCAAGTGATTTCAAGTGCTCTATCAAGA 60
QY 1695 GGAGATAGCAATTACAGTGGTGAATGGGAAGTCTTGGCCGGCATGATCTAATATCCA 1754
DB 61 GGAGATAGCAATTACAGTGGTGAATGGGAAGTCTTGGCCGGCATGATCTAATATCCA 120
QY 1755 AGTTGATGAAGTCAGAGGCTGTATATTTTGAAGGCCAACAGCTCCCTTTAGAGCA 1814
DB 121 AGTTGATGAAGTCAGAGGCTGTATATTTTGAAGGCCAACAGCTCCCTTTAGAGCA 180
QY 1815 TCACCTAGTGTAGTCAGTACGTAATCTCGAGAGGTGACAGGCTGACTACCGTGG 1874
DB 181 TCACCTAGTGTAGTCAGTACGTAATCTCGAGAGGTGACAGGCTGACTACCGTGG 240
QY 1875 CTACTCATCTTTGCTGCATCAGTCAGCAGCTGTGACTCTTTTAAGTAAGTATAGTAA 1934
DB 241 CTACTCATCTTTGCTGCATCAGTCAGCAGCTGTGACTCTTTTAAGTAAGTATAGTAA 300
QY 1935 CCAGAAGATCCACACTGTGTGTCCTTTTACAGCTATCAAGTCTGAAGATGACCCAAC 1994
DB 301 CCAGAAGATCCACACTGTGTGTCCTTTTACAGCTATCAAGTCTGAAGATGACCCAAC 360
QY 1995 TTGCAGAAACAAAGGAATTTTGGGCCACCACTTTTGGATTTTCAGCAGGTCTCTTCTGACTA 2054
DB 361 TTGCAGAAACAAAGGAATTTTGGGCCACCACTTTTGGATTTTCAGCAGGTCTCTTCTGACTA 420

QY 2055 TACTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGAATTACATTGATGGATGCT 2114
DB 421 TACTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGAATTACATTGATGGATGCT 480
QY 2115 CTACAAGCCTCATGATCTACAGCCTGGAAGAAATATCCTACTGCTGTTTATATATGG 2174
DB 481 CTACAAGCCTCATGATCTACAGCCTGGAAGAAATATCCTACTGCTGTTTATATATGG 540
QY 2175 TGGTCTCAGGTGAGTGGTGAATTAATCGGTTTAAAGAGAGTCAAGTATTTCCGCTTGA 2234
DB 541 TGGTCTCAGGTGAGTGGTGAATTAATCGGTTTAAAGAGAGTCAAGTATTTCCGCTTGA 600
QY 2235 TACCTAGCCTCTCTAGGTTATGTGGTTAGTGTAGTACACACAGGGGATCCTGTCCCG 2294
DB 601 TACCTAGCCTCTCTAGGTTATGTGGTTAGTGTAGTACACACAGGGGATCCTGTCCCG 660
QY 2295 AGGCTTTAAATTTGAAGGGCCCTTTAAATATATAAATGGTCAATAGAAAATTGACGATCA 2354
DB 661 AGGCTTTAAATTTGAAGGGCCCTTTAAATATATAAATGGTCAATAGAAAATTGACGATCA 720
QY 2355 GGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTTCAATGCTTAGATCGTGGG 2414
DB 721 GGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTTCAATGCTTAGATCGTGGG 780
QY 2415 CATCCAGCGCTGCTTATGGAGGATACCTTCCCTGATGGCATT 2459
DB 781 CATCCAGCGCTGCTTATGGAGGATACCTTCCCTGATGGCATT 825
RESULT 10
AF176779
LOCUS AF176779 1265 bp mRNA linear HTC 01-AUG-2003
DEFINITION Homo sapiens MSTP141 mRNA, complete cds.
ACCESSION AF176779
VERSION AF176779.1 GI:33338069
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 1265)
Hui,R.T., Liu,Y.Q., Wang,X.Y., Qin,B.M. and Sheng,H.
Homo sapiens normal aorta mRNA MST141
Unpublished
2 (bases 1 to 1265)
Hui,R.T., Liu,Y.Q., Wang,X.Y., Qin,B.M. and Sheng,H.
Direct Submission
Submitted (10-AUG-1999) Molecular Medicine Center for
Cardiovascular Disease, Cardiovascular Institute, CAMS & PUMC, 167,
Bei Li Shi Lu, Beijing 100037, P.R. China
FEATURES
Location/Qualifiers
1..1265
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="aorta"
268..741
/codon_start=1
/product="MSTP141"
/protein_id="AAQ13657.1"
/db_xref="GI:33338070"
/translation="MLYKPHDLQPGKYPVTLFIYGGFQVOLNNRPFKVKYFRLNLT
ASLGYVVVVIDNRGSGCHRLKFGKYMGOIEIDQVEGLQYLASRYDFIDLDVRG
IHGWSYGYLSLMLMQRSDIFRVAIAGAPVTLWIFYDTGYTERYMGHPDQEQG"
ORIGIN
Query Match 23.1%; Score 722; DB 4; Length 1265;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 722; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1861 CTGACTGACCGTGGCTACTCACATCTTCTGTCATCAGTACGACTGTGACTTCTTTATA 1920


```
Db 19 CTGACTGACGCTGGCTACATCTCTTCTGTCATCAGTCAGCACTGTGATCTTTTATA 78
Qy 1921 AGTAAGTATAGTAACAGAGAAGATCCACACTGTGTGTCCTTTTAAAGCTATCAAGTCCT 1980
Db 79 AGTAAGTATAGTAACAGAGAAGATCCACACTGTGTGTCCTTTTAAAGCTATCAAGTCCT 138
Qy 1981 GAAGATGACCCAACTTGCAGAAACAAAGAAATTTGGGCCACCATTTTGGATTGACAGGT 2040
Db 139 GAAGATGACCCAACTTGCAGAAACAAAGAAATTTGGGCCACCATTTTGGATTGACAGGT 198
Qy 2041 CCTCTTCTGACTATCTCTCCAGAAATTTTCTTTTGAAGTACTACTGATTTTACA 2100
Db 199 CCTCTTCTGACTATCTCTCCAGAAATTTTCTTTTGAAGTACTACTGATTTTACA 258
Qy 2101 TTGTATGGGATGCTCTACAGGCTCATGATCTACAGCTCGAAGCTGGAAGAAATATCTTACTGTG 2160
Db 259 TTGTATGGGATGCTCTACAGGCTCATGATCTACAGCTCGAAGCTGGAAGAAATATCTTACTGTG 318
Qy 2161 CTGTTCATATATATGTTGGTCTCTCAGGTGCAAGTTGGTGAATATCGGTTTAAAGGAGTCAAG 2220
Db 319 CTGTTCATATATGTTGGTCTCTCAGGTGCAAGTTGGTGAATATCGGTTTAAAGGAGTCAAG 378
Qy 2221 TATTTCCGTTGAATACCTTAGCTCTCTAGGTTATGTTGTTAGTGTATGATGATAGACACAGG 2280
Db 379 TATTTCCGTTGAATACCTTAGCTCTCTAGGTTATGTTGTTAGTGTATGATGATAGACACAGG 438
Qy 2281 GATCTCTGACCGAGGCTTAAATTTGAAGGCGCTTAAATATAAATGGGTCAATA 2340
Db 439 GATCTCTGACCGAGGCTTAAATTTGAAGGCGCTTAAATATAAATGGGTCAATA 498
Qy 2341 GAAATTGACATCAGGTGGAAGACTCCAAATATCTAGCTTCTCGATATGATTTTCATTGAC 2400
Db 499 GAAATTGACATCAGGTGGAAGACTCCAAATATCTAGCTTCTCGATATGATTTTCATTGAC 558
Qy 2401 TTAGATCGTGTGGGCAATCCAGGCTGCTCTATGAGGATACCTCTCCCTGATGGCATTA 2460
Db 559 TTAGATCGTGTGGGCAATCCAGGCTGCTCTATGAGGATACCTCTCCCTGATGGCATTA 618
Qy 2461 ATGACAGGTCAGATATCTTACAGGTTGCTATTTGCTGGGGCCCCAGTCACCTCTGTGATC 2520
Db 619 ATGACAGGTCAGATATCTTACAGGTTGCTATTTGCTGGGGCCCCAGTCACCTCTGTGATC 678
Qy 2521 TTCTATGATACAGATACAGGAACCTTATATGTTGGTCACTCCCTGACAGATGAACAGGC 2580
Db 679 TTCTATGATACAGATACAGGAACCTTATATGTTGGTCACTCCCTGACAGATGAACAGGC 738
Qy 2581 TA 2582
Db 739 TA 740

RESULT 11
AL043338 735 bp mRNA linear EST 04-SEP-2003
LOCUS DKEZp43400723.r1.434 (synonym: htes3) Homo sapiens cdna clone
DEFINITION DKEZp43400723-5', mRNA sequence.
ACCESSION AL043338
VERSION AL043338.1 GI:5422728
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 735)
REFERENCE Blum, H., Bauersachs, S., Mewes, H.W., Gassenhuber, J. and Wiemann, S.
AUTHORS EST (Blum, et al.)
TITLE Unpublished (1999)
JOURNAL Contact: MIPS
COMMENT MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
```

Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de; sequenced by LMU (Ludwig Maximilians University, Munich/Germany) within the cDNA sequencing consortium of the German Genome Project.
No s1 sequence available.
This clone (DKFZp43400723) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
Location/Qualifiers
1..735
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp43400723"
/tissue_type="testis"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="434 (synonym: htes3)"
/note="Vector: pSport1; Site_1: NotI; Site_2: SalI"

FEATURES
source

ORIGIN

```
Query Match 22.7%; Score 708; DB 1; Length 735;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 708; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 2274 CAACAGGGGATCCTGTCCAGGCGCTTAAATTTGAAGCGCCTTAAATAATAAATGGG 2333
Db 1 CAACAGGGGATCCTGTCCAGGCGCTTAAATTTGAAGCGCCTTAAATAATAAATGGG 60
Qy 2334 TCAATAAGAAATGACGATCAGGTGAAGGATCCAAATATCTAGCTTCTCGATGATTT 2393
Db 61 TCAATAAGAAATGACGATCAGGTGAAGGATCCAAATATCTAGCTTCTCGATGATTT 120
Qy 2394 CATTGACTTAGATCGTGTGGGCATCCAGGCTGGTCTTATGGAGTACCTTCCCTGAT 2453
Db 121 CATTGACTTAGATCGTGTGGGCATCCAGGCTGGTCTTATGGAGTACCTTCCCTGAT 180
Qy 2454 GGCATTAATCCAGGTCAGATATCTTACGGGTGCTATTGCTGGGGCCCCAGTCACCTCT 2513
Db 181 GGCATTAATCCAGGTCAGATATCTTACGGGTGCTATTGCTGGGGCCCCAGTCACCTCT 240
Qy 2514 GTGGATCTTCTATGATACAGGATACACGGAAGCTTATATGGGTCACTCCCTGACGAATGA 2573
Db 241 GTGGATCTTCTATGATACAGGATACACGGAAGCTTATATGGGTCACTCCCTGACGAATGA 300
Qy 2574 ACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAAAGTTCCCTCTGACACAAA 2633
Db 301 ACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAAAGTTCCCTCTGACACAAA 360
Qy 2634 TCGTTTACTGCTCTTACATGTTTCTGATGAGATGTCATTTTGACATACCATGAT 2693
Db 361 TCGTTTACTGCTCTTACATGTTTCTGATGAGATGTCATTTTGACATACCATGAT 420
Qy 2694 ATTACTCAGTTTTTTAGTGGGCTGGAAGCCATATGATTTTACAGATCTTCTCCAGGA 2753
Db 421 ATTACTCAGTTTTTTAGTGGGCTGGAAGCCATATGATTTTACAGATCTTCTCCAGGA 480
Qy 2754 GAGACAGCATTAAGAGTTCCTCAATCGGAGAACATTTATGAATGTCATCTTTTGCATTA 2813
Db 481 GAGACAGCATTAAGAGTTCCTCAATCGGAGAACATTTATGAATGTCATCTTTTGCATTA 540
Qy 2814 CTTTCAAGAAAACTTGGATCAGTATTTGCTCTTAAAGTGATATATTTTGACCTGT 2873
Db 541 CTTTCAAGAAAACTTGGATCAGTATTTGCTCTTAAAGTGATATATTTTGACCTGT 600
Qy 2874 GTAGAACTCTCTGTATACACTGGCTATTTAAACCAATCAGGAGGTTTAAATCAACAGAAA 2933
Db 601 GTAGAACTCTCTGTATACACTGGCTATTTAAACCAATCAGGAGGTTTAAATCAACAGAAA 660
Qy 2934 ACACAGAAATGATCATCATTTTGTATCTGCTGATCTGATCTTAACTTACT 2981
Db 661 ACACAGAAATGATCATCATTTTGTATCTGCTGATCTGATCTTAACTTACT 708
```

RESULT 12	AL040398	AL040398	753 bp	mRNA	linear	EST 04-SEP-2003
LOCUS	DKFZp344A0714	r1 434 (synonym: htes3)	Homo sapiens	cDNA clone		
DEFINITION	DKFZp344A0714 5', mRNA sequence.					
ACCESSION	AL040398					
VERSION	AL040398.1	GI:5409350				
KEYWORDS	EST.					
SOURCE	Homo sapiens (human)					
ORGANISM	Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.					
REFERENCE	1 (bases 1 to 753)					
AUTHORS	Koehrer, K., Beyer, A., Mewes, H.W., Gassenhuber, J. and Wiemann, S.					
JOURNAL	EST (Koehrer, et al.)					
COMMENT	Unpublished (1999)					
	Contact: MIPS					
	MIPS					
	Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany					
	This is the 5' sequence of the clone insert					
	Clone from S. Wiemann, Molecular Genome Analysis, German Cancer					
	Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;					
	sequenced by BMFZ (Biomedical Research Center at the Charite,					
	Berlin/Germany) within the cDNA sequencing consortium of the German					
	Genome Project.					
	No sl sequence available.					
	This clone (DKFZp344A0714) is available at the RZPD in Berlin.					
	Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059					
	Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.					
FEATURES	Location/Qualifiers					
source	1..753					
	/organism="Homo sapiens"					
	/mol_type="mRNA"					
	/db_xref="taxon:9606"					
	/clone="DKFZp344A0714"					
	/tissue_type="testis"					
	/dev_stage="adult"					
	/lab_host="ph10B"					
	/clone_lib="434 (synonym: htes3)"					
	/note="Vector: pSport1; Site_1: NotI; Site_2: SalI"					
ORIGIN						
Query Match	22.4%; Score 700; DB 1; Length 753;					
Best Local Similarity	99.9%; Pred. No. 0;					
Matches	750; Conservative 0; Mismatches 1; Indels 0; Gaps 0;					
Qy	1609 ATTACATCTATTTTAAAGGAAAGCAAAATATAAAGCATCCAGTGTGGGCTGCCTGCTCCA	1668				
Db	1 ATTACATCTATTTTAAAGGAAAGCAAAATATAAAGCATCCAGTGTGGGCTGCCTGCTCCA	60				
Qy	1669 AGTGATTTCAAGTGTCTATCAAGAGGAGATAGCAATTACCAAGTGTGAATGGGAAGTT	1728				
Db	61 AGTGATTTCAAGTGTCTATCAAGAGGAGATAGCAATTACCAAGTGTGAATGGGAAGTT	120				
Qy	1729 CTTGGCCGGCATGGATCTAATATCAAGTGTCAAGTGTCAAGAGGCTGGTATATTTTGA	1788				
Db	121 CTTGGCCGGCATGGATCTAATATCAAGTGTCAAGTGTCAAGAGGCTGGTATATTTTGA	180				
Qy	1789 GGCACAAAGATCCCTTTAGAGCATCACCTGTACGTACGTACGTACGTACGTACGTACGT	1848				
Db	181 GGCACAAAGATCCCTTTAGAGCATCACCTGTACGTACGTACGTACGTACGTACGTACGT	240				
Qy	1849 GAGTCACAAAGCTGACTGACCGTGGTACTACATCTTGTGCTGCTGCTGCTGCTGCTG	1908				
Db	241 GAGTCACAAAGCTGACTGACCGTGGTACTACATCTTGTGCTGCTGCTGCTGCTGCTG	300				
Qy	1909 GACTCTTTTAAAGTATAGTATACGAGGAATCCACATCTGTGTGCTTCCCTTACAG	1969				
Db	301 GACTCTTTTAAAGTATAGTATACGAGGAATCCACATCTGTGTGCTTCCCTTACAG	360				
Qy	1969 CTATCAAGTCTCTGAAGATGACCCAACTTGCAGAAACAAAGGAATTTTGGCCACCAATTTTG	2028				
Db	361 CTATCAAGTCTCTGAAGATGACCCAACTTGCAGAAACAAAGGAATTTTGGCCACCAATTTTG	420				
Qy	2029 GATTACAGAGGTCTCTCTGACTACTACTCTCCAGAAATTTTCTCTTTTGAAGTACT	2088				
Db	421 GATTACAGAGGTCTCTCTGACTACTACTCTCCAGAAATTTTCTCTTTTGAAGTACT	480				
Qy	2089 ACTGGATTTACATGTTATGGGATGCTCTACAAGCCTCATGATCTACAGCTGGAAAGAAA	2148				
Db	481 ACTGGATTTACATGTTATGGGATGCTCTACAAGCCTCATGATCTACAGCTGGAAAGAAA	540				
Qy	2149 TATCCTACTGTGCTGTTTATATATGTTGCTCTCAGGTGCAGTGTGGTGAATATCGGTTT	2208				
Db	541 TATCCTACTGTGCTGTTTATATATGTTGCTCTCAGGTGCAGTGTGGTGAATATCGGTTT	600				
Qy	2209 AAAGGAGTCAAGTATTTCCGCTTTGAATACCCTAGCCTCTCTAGGTATGTGGTGTAGTG	2268				
Db	601 AAAGGAGTCAAGTATTTCCGCTTTGAATACCCTAGCCTCTCTAGGTATGTGGTGTAGTG	660				
Qy	2269 ATAGACAAACAGGGGATCCTGTCAACCGAGGCTTTAAATTTGAAGGCGCTTTAAATATAAA	2328				
Db	661 ATAGACAAACAGGGGATCCTGTCAACCGAGGCTTTAAATTTGAAGGCGCTTTAAATATAAA	720				
Qy	2329 ATGGGTCAAAATACAAATTTGACGATCAGGTGG	2359				
Db	721 ATGGGTCAAAATACAAATTTGACGATCAGGTGG	751				
RESULT 13	CR998849	RZPD no. 9016	Homo sapiens	cDNA clone	linear	EST 28-JUN-2005
LOCUS	CR998849					
DEFINITION	CR998849 RZPD no. 9016 Homo sapiens mRNA sequence.					
ACCESSION	CR998849					
VERSION	CR998849.1	GI:68292734				
KEYWORDS	EST.					
SOURCE	Homo sapiens (human)					
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.					
REFERENCE	1 (bases 1 to 835)					
AUTHORS	Heil, O., Ebert, L., Hennig, S., Henze, S., Radelof, U., Schneider, D. and Korn, B.					
TITLE	Human T-Lymphocytes library					
JOURNAL	Unpublished (2005)					
COMMENT	Contact: Inge Airlart					
	RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH					
	Heubnerweg 6, D-14059 Berlin, Germany					
	Email: www.rzpd.de					
	RZPD: RZPDp9016K2027.					
	RZPDLIB: (Human T-Lymphocytes) RZPD LIB No. 9016					
	http://www.rzpd.de/cgi-bin/products/set.cgi?libNo=9016					
	Contact: Inge Airlart					
	RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH					
	Heubnerweg 6, D-14059 Berlin, Germany					
	Tel: +49 30 32639 100					
	Fax: +49 30 32639 111					
	www.rzpd.de					
	This clone is available from RZPD;					
	contact RZPD (product- support@rzpd.de) for further information.					
	Primer name: q3_4, Primer sequence: CGGATAACAATTCACACAG.					
FEATURES	Location/Qualifiers					
source	1..835					
	/organism="Homo sapiens"					
	/mol_type="mRNA"					
	/db_xref="taxon:9606"					
	/clone="RZPDp9016K2027"					
	/tissue_type="T-Lymphocytes"					
	/dev_stage="adult"					
	/lab_host="DH10B"					

/clone_lib="RZPD no.9016"
/note="vector: pQE80LSN_cloned; Site_1: Sali; Site_2:
NotI; vector:
http://www.rzpd.de/info/vectors/pQE80LSN_cloned.p1c.shtml
; 1st strand cDNA was prepared from mRNA obtained from
human T-lymphocytes with a NotI - oligo(dT) primer [5'
GACTAGTCTAGATCGGCGCGCCGCTTTTCTTTTCTTTT 3']
Double-stranded cDNA was ligated to Sali adaptors,
digested with NotI and cloned into the NotI and Sali sites
of the pQE80LSN_cloned vector"

ORIGIN

Query Match 22.1%; Score 690; DB 7; Length 835;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 740; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 78 GTCAACGGCGCGCGCGCGAGGAGCCACTGCAACCGAGCCGAGTGGAGCGCGCGAG 137
Db 1 GTCAACGGCGCGCGCGCGAGGAGGAGCCACTGCAACCGAGCCGAGTGGAGCGCGCGAG 60

QY 138 CATGAAGCGCGAGCGCGCGCTCATAGCGACGTCGGGACGCTCGGCGCGCGCGCGGG 197
Db 61 CATGAAGCGCGAGCGCGCGCTCATAGCGACGTCGGGACGCTCGGCGCGCGCGCGGG 120

QY 198 GGAAGGAATGCAACATGCGCAGCAGCAATGGAAACAGACAGCTGGGTGTGAGATATT 257
Db 121 GGAAGGAATGCAACATGCGCAGCAGCAATGGAAACAGACAGCTGGGTGTGAGATATT 180

QY 258 TGAACATGCGGACTGTGAGGAGAAATTTGAATCACAGGATGGCTTAAATGGAGCCCTTT 317
Db 181 TGAACATGCGGACTGTGAGGAGAAATTTGAATCACAGGATGGCTTAAATGGAGCCCTTT 240

QY 318 TTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAGTCGTTCGCCGATACCAAGAAATA 377
Db 241 TTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAGTCGTTCGCCGATACCAAGAAATA 300

QY 378 TCATGGCTTACATGATGCTAAGGCACCAACATGATTTTCAATGTTTGAAGAGGAATGATCC 437
Db 301 TCATGGCTTACATGATGCTAAGGCACCAACATGATTTTCAATGTTTGAAGAGGAATGATCC 360

QY 438 AGATGGACCTCATTCAGACAGAAATCTATTACCTTGGCCATGCTGTGTGAGAACAGAGAAA 497
Db 361 AGATGGACCTCATTCAGACAGAAATCTATTACCTTGGCCATGCTGTGTGAGAACAGAGAAA 420

QY 498 TACACTGTTTATCTGAAATTCCTAAACTATCATAGACAGCAGCTTTAAATGCTCTC 557
Db 421 TACACTGTTTATCTGAAATTCCTAAACTATCATAGACAGCAGCTTTAAATGCTCTC 480

QY 558 TTGAAGCCTCTTTTGGATCTTTTCAGGCAACACTGGACTATGGAATGTTTCTCGAGA 617
Db 481 TTGAAGCCTCTTTTGGATCTTTTCAGGCAACACTGGACTATGGAATGTTTCTCGAGA 540

QY 618 AGAAGAACTATTAAAGAGAAAGAAACGCAATGGAACAGTCGGAATGCTTCTTACGATTA 677
Db 541 AGAAGAACTATTAAAGAGAAAGAAACGCAATGGAACAGTCGGAATGCTTCTTACGATTA 600

QY 678 TCACCAAGGAGTGGAAATTTCTGTTTCAAGCGGTAGTGGAAATTTATCATCGTAAAGA 737
Db 601 TCACCAAGGAGTGGAAATTTCTGTTTCAAGCGGTAGTGGAAATTTATCATCGTAAAGA 660

QY 738 TGGAGGGCCACAGAGTTTACGCAACACTTTAAGGCCCAATCTAGTGGAAACTAGTTG 797
Db 661 TGGAGGGCCACAGAGTTTACGCAACACTTTAAGGCCCAATCTAGTGGAAACTAGTTG 720

QY 798 TCCCAACATACGGATGATCC 818
Db 721 TCCCAACATACGGATGATCC 741

RESULT 14

CN427229 742 bp mRNA linear EST 16-MAY-2004
LOCUS CN427229
DEFINITION 17000600259975 GRN_PRENU Homo sapiens cDNA 5', mRNA sequence.

ACCESSION CN427229
VERSION CN427229.1 GI:47414823
KEYWORDS EST;
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 742)
AUTHORS Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G.J.,
Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M.S., Mandalam, R.,
Lebkowski, J. and Stanton, L.W.
TITLE Transcriptome characterization elucidates signaling networks that
control human ES cell growth and differentiation
JOURNAL Nat. Biotechnol. 22 (6), 707-716 (2004)
PUBMED 15146197
COMMENT Contact: Brandenberger R
Regenerative Medicine
Geron Corporation
230 Constitution Drive, Menlo Park, CA 94025, USA
Tel: 650 473 8658
Fax: 650 473 7760
Email: rbrandenberger@geron.com
Insert Length: 742 Std Error: 0.00.
FEATURES
source
1..742
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="embryonic stem cell, retinoic acid and
mitogen-treated hES cell line H7"
/clone_lib="GRN_PRENU"
/note="oligo dT primed, full-length enriched cDNA library
from hES cell line H7 (p29) maintained in feeder-free
conditions. Embryoid bodies were generated in the presence
of all-trans retinoic and mitogens."

ORIGIN
Query Match 22.0%; Score 687; DB 7; Length 742;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 737; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 552 GCTCTCTTGAAGCCTCTTTTGGATCTTTTTCAGGCAACACTGGACTATGGAATGATTC 611
Db 5 GCTCTCTTGAAGCCTCTTTTGGATCTTTTTCAGGCAACACTGGACTATGGAATGATTC 64

QY 612 TCGAGAGAGAACTATTAAAGAGAAAGAAACCCATTTGGAACAGTCGGAATGCTTCTTA 671
Db 65 TCGAGAGAGAACTATTAAAGAGAGAAAGAAACCCATTTGGAACAGTCGGAATGCTTCTTA 124

QY 672 CGATTATCAACAGGAAGTGGAAACATTTCTGTTTCAAGCGGTAGTGGAAATTTATCAGT 731
Db 125 CGATTATCAACAGGAAGTGGAAACATTTCTGTTTCAAGCGGTAGTGGAAATTTATCAGT 184

QY 732 AAAAGATGGAGGCGCACAGGATTTACGGAACAACTTTAAGGCCCAATCTAGTGGAAAC 791
Db 185 AAAAGATGGAGGCGCACAGGATTTACGGAACAACTTTAAGGCCCAATCTAGTGGAAAC 244

QY 792 TAGTTGTCCCAACATACGGATGGATCCAAATTTATGCCCGGTGATCCAGACTGGATTCG 851
Db 245 TAGTTGTCCCAACATACGGATGGATCCAAATTTATGCCCGGTGATCCAGACTGGATTCG 304

QY 852 TTTTATACATAGCAACGATATTTGGATATCTAACTCGTAACCAAGAGAAAGAGACT 911
Db 305 TTTTATACATAGCAACGATATTTGGATATCTAACTCGTAACCAAGAGAAAGAGACT 364

QY 912 CACTTATGTGCAATGAGCTAGCCAACTGGGAAGAGATGCCAGATCAGCTGGAGTCGC 971
Db 365 CACTTATGTGCAATGAGCTAGCCAACTGGGAAGAGATGCCAGATCAGCTGGAGTCGC 424

QY 972 TACCTTTGTTCTCCAAAGAAATTTGATAGATATTTCTGGCTATTTGGTGGTGTCCAAAAGC 1031
Db 425 TACCTTTGTTCTCCAAAGAAATTTGATAGATATTTCTGGCTATTTGGTGGTGTCCAAAAGC 484

```
QY 1032 TCACAACTCCAGTGGTGGTAAATTTCTAGAAATCTATATGAAGAAATGATGAATC 1091
Db 485 TGAACAACTCCAGTGGTGGTAAATTTCTAGAAATCTATATGAAGAAATGATGAATC 544
QY 1092 TGAGGTGGAATTAATTCATGTTACATCCCTATGTTGGAACAAGAGGGGAGATTCATT 1151
Db 545 TGAGGTGGAATTAATTCATGTTACATCCCTATGTTGGAACAAGAGGGGAGATTCATT 604
QY 1152 CCGTTATCCTAAACAGGTACAGCAATCCTAAAGTCACTTTTAAGATGTCAGAAATAAT 1211
Db 605 CCGTTATCCTAAACAGGTACAGCAATCCTAAAGTCACTTTTAAGATGTCAGAAATAAT 664
QY 1212 GATTGATCCTGAAGAGGATCATAGATGTCATAGATAAGGAATCAATTCACCTTTTGA 1271
Db 665 GATTGATCCTGAAGAGGATCATAGATGTCATAGATAAGGAATCAATTCACCTTTTGA 724
QY 1272 GATTCTATTTGAAGGAGT 1289
Db 725 GATTCTATTTGAAGGAGT 742

RESULT 15
BX390898
LOCUS BX390898 Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED
DEFINITION Homo sapiens cDNA clone CS0DL005YD02 5-PRIME, mRNA sequence.
ACCESSION BX390898
VERSION BX390898.2 GI:46844267
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
1 (bases 1 to 914)
Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On May 13, 2003 this sequence version replaced gi:30607432.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
7542.r
For More information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0BAG0532D07_CS05044_1&c=7542.r
.
FEATURES
Source Location/Qualifiers
1..914
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DL005YD02"
/call_type="B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"
/call_lib="Homo sapiens B CELLS (RAMOS CELL LINE) COT
25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V sites of the pCMVSPORT 6
vector. Library was normalized."
ORIGIN
Query Match 22.0%; Score 686; DB 5; Length 914;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 736; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 739 CGAGGCCACACAGGATTTAGCGAACAAACCTTTTAAGGCCCAATCTACTCGAAACTAGTTGT 798
Db 73 CGAGGCCACACAGGATTTAGCGAACAAACCTTTTAAGGCCCAATCTACTCGAAACTAGTTGT 132
QY 799 CCCAACATACAGGATGGATCCAAAATTTATGCCCGCTGTATCCAGACTGGATTCGTTTTATA 858
Db 133 CCCAACATACAGGATGGATCCAAAATTTATGCCCGCTGTATCCAGACTGGATTCGTTTTATA 192
QY 859 CATAGCAACGATATTTGGATATCTAAATCATCGTAAACAGAGAAAGAGAGACTCACTTAT 918
Db 193 CATAGCAACGATATTTGGATATCTAAATCATCGTAAACAGAGAAAGAGAGACTCACTTAT 252
QY 919 GTGCACAAATGAGCTAGCAGCAACATGGAAGAGTGCAGATCAGCTGGAGTGCCTACCTTT 978
Db 253 GTGCACAAATGAGCTAGCAGCAACATGGAAGAGTGCAGATCAGCTGGAGTGCCTACCTTT 312
QY 979 GTTCTCTCAAGAAGAAATTTGATAGATATCTGGCTATTTGGTGGTGTCCAAAAGCTGAAACA 1038
Db 313 GTTCTCTCAAGAAGAAATTTGATAGATATCTGGCTATTTGGTGGTGTCCAAAAGCTGAAACA 372
QY 1039 ACTCCCAAGTGGTGAATAATTTCTAGAAATTTCTATATGAAGAAATATGATGAATCTGAGTG 1098
Db 373 ACTCCCAAGTGGTGAATAATTTCTAGAAATTTCTATATGAAGAAATATGATGAATCTGAGTG 432
QY 1099 GAAATTTATCATGTTACATCCCTATGTTGGAACAAGAGGGGAGATTCATTCCGTTAT 1158
Db 433 GAAATTTATCATGTTACATCCCTATGTTGGAACAAGAGGGGAGATTCATTCCGTTAT 492
QY 1159 CCTAARACAGGTACAGCAATCCTAAAGTCACTTTTAAAGTGTCAAGAAATATGATGAT 1218
Db 493 CCTAARACAGGTACAGCAATCCTAAAGTCACTTTTAAAGTGTCAAGAAATATGATGAT 552
QY 1219 GCTGAAGAAAGGATCATAGATGTCATAGATAAGGAATCAATTCACCTTTTGAGATTTCTA 1278
Db 553 GCTGAAGAAAGGATCATAGATGTCATAGATAAGGAATCAATTCACCTTTTGAGATTTCTA 612
QY 1279 TTTGAAGAGTTGAATATATTTGCCAGAGCTGGATGGACTCTCTGAGGGAATAATGCTTGG 1338
Db 613 TTTGAAGAGTTGAATATATTTGCCAGAGCTGGATGGACTCTCTGAGGGAATAATGCTTGG 672
QY 1339 TCCATCTCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAATTA 1398
Db 673 TCCATCTCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAATTA 732
QY 1399 TTTATCCCAAGTGAAGATGATGTTATGGAAGGAGAGAGACTCATTGAGTGCCTGAT 1458
Db 733 TTTATCCCAAGTGAAGATGATGTTATGGAAGGAGAGAGACTCATTGAGTGCCTGAT 792
QY 1459 TCTGTGACGCCACTAAT 1475
Db 793 TCTGTGACGCCACTAAT 809

Search completed: May 4, 2006, 12:52:35
Job time : 11802 secs
```


Db 241 CTGGGTGTTGAGATATTTGAAACTGGGACTGTGAGGAGAAATTTGAATCACAGGATCGG 300
Qy 301 CCTAAATTTGGAGCCCTTTTATGTTTGGACGGTATTCCTGGAGTCAGCTTAAAAAAGCTGCTT 360
Db 301 CCTAAATTTGGAGCCCTTTTATGTTTGGACGGTATTCCTGGAGTCAGCTTAAAAAAGCTGCTT 360
Qy 361 GCCGATACCAAGAAATATCATGCTACATGATGGCTTAAGGCACACATGATTTTCATGTTT 420
Db 361 GCCGATACCAAGAAATATCATGCTACATGATGGCTTAAGGCACACATGATTTTCATGTTT 420
Qy 421 GTGAAGAGGAATGATCCAGATGACCTCAATCAGACAGAACTTATACCTTGGCCATGCT 480
Db 421 GTGAAGAGGAATGATCCAGATGACCTCAATCAGACAGAACTTATACCTTGGCCATGCT 480
Qy 481 GGTGAGAACAGAGAAATACACTGTTTATCTGAAATTCGCAAACTATCAATPAGAGCA 540
Db 481 GGTGAGAACAGAGAAATACACTGTTTATCTGAAATTCGCAAACTATCAATPAGAGCA 540
Qy 541 CGAGTCTTAATGCTCTCTTGGAAAGCCTCTTTTGGATCTTTTCAGGCAACACTGGACTAT 600
Db 541 CGAGTCTTAATGCTCTCTTGGAAAGCCTCTTTTGGATCTTTTCAGGCAACACTGGACTAT 600
Qy 601 GGAATGATTTCTCGAGAAGAGAACTATTAAGAGAAAGAAACGCATTTGGAACAGTCGGA 660
Db 601 GGAATGATTTCTCGAGAAGAGAACTATTAAGAGAAAGAAACGCATTTGGAACAGTCGGA 660
Qy 661 ATTGCTTTCTTACGATTTATCACCAAGGAAGTGGAACATTTCTGTTTCAAGCCGGTAGTGA 720
Db 661 ATTGCTTTCTTACGATTTATCACCAAGGAAGTGGAACATTTCTGTTTCAAGCCGGTAGTGA 720
Qy 721 ATTATCACGTAAGAGATGGAGGCCACAAAGGATTTAGCAACAACTTTAAAGGCCCAAT 780
Db 721 ATTATCACGTAAGAGATGGAGGCCACAAAGGATTTAGCAACAACTTTAAAGGCCCAAT 780
Qy 781 CTAGTGGAACTAGTTTGTCCCAACATACGGATGATCCAAATTTATGCTCCCGCTGATCCA 840
Db 781 CTAGTGGAACTAGTTTGTCCCAACATACGGATGATCCAAATTTATGCTCCCGCTGATCCA 840
Qy 841 GACTGGATTTGCTTTATACATAGCAACGATATTTGGATATCTAATCATCGTAACACAGAA 900
Db 841 GACTGGATTTGCTTTATACATAGCAACGATATTTGGATATCTAATCATCGTAACACAGAA 900
Qy 901 GAAAGGAGACTCACTTATGTGCACAAATGAGCTAGCCTAACCAATGGAAGAAGATGCCAGATCA 960
Db 901 GAAAGGAGACTCACTTATGTGCACAAATGAGCTAGCCTAACCAATGGAAGAAGATGCCAGATCA 960
Qy 961 GCTGGAGTCGCTACCTTTGTTCTCAAGAGAAATTTGATAGATAATTTCTGGCTATTGGTGG 1020
Db 961 GCTGGAGTCGCTACCTTTGTTCTCAAGAGAAATTTGATAGATAATTTCTGGCTATTGGTGG 1020
Qy 1021 TGTCCTAAAGCTGAAACAACTCCAGTGGTGGTAAAAATTTCTTAGAATTTCTATATGAAGAA 1080
Db 1021 TGTCCTAAAGCTGAAACAACTCCAGTGGTGGTAAAAATTTCTTAGAATTTCTATATGAAGAA 1080
Qy 1081 AATGATGAATCTGAGGTGGAATTTATCATGTATACATCCCTATGTTGGAACAAGGAGG 1140
Db 1081 AATGATGAATCTGAGGTGGAATTTATCATGTATACATCCCTATGTTGGAACAAGGAGG 1140
Qy 1141 GCAGATTCATTCGTTTATCCTTAAACAGGTACACCAATCCTTAAGTCATCTTTTAAGATG 1200
Db 1141 GCAGATTCATTCGTTTATCCTTAAACAGGTACACCAATCCTTAAGTCATCTTTTAAGATG 1200
Qy 1201 TCAGAAATATCATGATGCTGMAAGGAGGATCATAGTCTCATAGATAGGAACTAATTT 1260
Db 1201 TCAGAAATATCATGATGCTGMAAGGAGGATCATAGTCTCATAGATAGGAACTAATTT 1260
Qy 1261 CAACCTTTGAGATTTCTATTTGAAGGAGTTGATATTTTCCAGAGCTGGATGGACTCT 1320
Db 1261 CAACCTTTGAGATTTCTATTTGAAGGAGTTGATATTTTCCAGAGCTGGATGGACTCT 1320
Qy 1321 GAGGAAATATGCTTTGGTCCATCCTACTAGATCGCTCCAGACTCGCCTACAGATAGTG 1380
Db 1321 GAGGAAATATGCTTTGGTCCATCCTACTAGATCGCTCCAGACTCGCCTACAGATAGTG 1380

Qy 1381 TTGATCTCACCTGAAATTTATTTATCCAGTAGAAGATGATGTTTATGGAAGGCAGAGACTC 1440
Db 1381 TTGATCTCACCTGAAATTTATTTATCCAGTAGAAGATGATGTTTATGGAAGGCAGAGACTC 1440
Qy 1441 ATTGAGTCAGTGCCCTGATTTCTGAGACGCCACTAAATTTATCTATGAAGAAACAACAGACATC 1500
Db 1441 ATTGAGTCAGTGCCCTGATTTCTGAGACGCCACTAAATTTATCTATGAAGAAACAACAGACATC 1500
Qy 1501 TGGATAAATATCCATGACATCTTTTTCATGTTTTTCCCAAGTCACGAAGAGGAAATTCAG 1560
Db 1501 TGGATAAATATCCATGACATCTTTTTCATGTTTTTCCCAAGTCACGAAGAGGAAATTCAG 1560
Qy 1561 TTTATTTTTCCTCTCTGAAATGCAAAAACAGGTTTCCGTCATTTTATACAAAATTTACATCTATT 1620
Db 1561 TTTATTTTTCCTCTCTGAAATGCAAAAACAGGTTTCCGTCATTTTATACAAAATTTACATCTATT 1620
Qy 1621 TTAAGGAAAGCAAAATATAAAGATCCAGTGGTGGCTGCCCTGCTCCAAAGTGAATTCAG 1680
Db 1621 TTAAGGAAAGCAAAATATAAAGATCCAGTGGTGGCTGCCCTGCTCCAAAGTGAATTCAG 1680
Qy 1681 TGTCTTATCAAGAGAGATAGCAATTAACAGTGGTGAATGGAGTTTCTTGGCCGGCAT 1740
Db 1681 TGTCTTATCAAGAGAGATAGCAATTAACAGTGGTGAATGGAGTTTCTTGGCCGGCAT 1740
Qy 1741 GGAATCTAATATCCAAGTTGATGAACTCAGAGGCTGGTATATTTTGAAGGCACCAAGAC 1800
Db 1741 GGAATCTAATATCCAAGTTGATGAACTCAGAGGCTGGTATATTTTGAAGGCACCAAGAC 1800
Qy 1801 TCCGCTTTAGAGATCACCTGTACGTAGTACGTAAATCTCTGAGAGGTGACAAGG 1860
Db 1801 TCCGCTTTAGAGATCACCTGTACGTAGTACGTAAATCTCTGAGAGGTGACAAGG 1860
Qy 1861 CTGACTGACCGTGGCTACTCAATTTCTGCTGCTGCACTAGTCAGCTGCTGCTTTTATA 1920
Db 1861 CTGACTGACCGTGGCTACTCAATTTCTGCTGCTGCACTAGTCAGCTGCTGCTTTTATA 1920
Qy 1921 AGTAAGTATAGTAAACAGAGAAATCCACACTGTGTGCTCCCTTTTACAAGCTATCAAGTCT 1980
Db 1921 AGTAAGTATAGTAAACAGAGAAATCCACACTGTGTGCTCCCTTTTACAAGCTATCAAGTCT 1980
Qy 1981 GAAGATGACCCAACTTGGCAAAACAAAGGAAATTTTGGGCCACCAATTTTGGATTCAGCAGGT 2040
Db 1981 GAAGATGACCCAACTTGGCAAAACAAAGGAAATTTTGGGCCACCAATTTTGGATTCAGCAGGT 2040
Qy 2041 CCTCTTCTGACTATACCTCCAGAAATTTCTCTTTGAAAGTACTACTGGATTTACA 2100
Db 2041 CCTCTTCTGACTATACCTCCAGAAATTTCTCTTTGAAAGTACTACTGGATTTACA 2100
Qy 2101 TTGTATGGGATGCTCTACAAGCCTCATGATCTCAGCTGCAAGAGAAATATCCTACTGTG 2160
Db 2101 TTGTATGGGATGCTCTACAAGCCTCATGATCTCAGCTGCAAGAGAAATATCCTACTGTG 2160
Qy 2161 CTGTTTCATATATGTTGTTCTCAGGTGCGATTTGGTGAATTAATCGGTTTAAAGAGTCAAG 2220
Db 2161 CTGTTTCATATATGTTGTTCTCAGGTGCGATTTGGTGAATTAATCGGTTTAAAGAGTCAAG 2220
Qy 2221 TATTTCCGTTGAATACCTTAGCTCTTAGTTTATGTTGTTAGTGATAGACAACAGG 2280
Db 2221 TATTTCCGTTGAATACCTTAGCTCTTAGTTTATGTTGTTAGTGATAGACAACAGG 2280
Qy 2281 GGAATCTGTCAACGAGGCTTAAATTTGAAGGCCCTTTTAAATATATAAATTTGGTCAAAATA 2340
Db 2281 GGAATCTGTCAACGAGGCTTAAATTTGAAGGCCCTTTTAAATATATAAATTTGGTCAAAATA 2340
Qy 2341 GAAATTTGACGATCAGGTGGAAGGACTCAATATCTAGCTTCTCGATATGATTTTATTGAC 2400
Db 2341 GAAATTTGACGATCAGGTGGAAGGACTCAATATCTAGCTTCTCGATATGATTTTATTGAC 2400
Qy 2401 TTAGATCGTGTGGGCATCCAGGCTGGTCTTATGGAGGATACCTCTCCCTGATGGCATTA 2460
Db 2401 TTAGATCGTGTGGGCATCCAGGCTGGTCTTATGGAGGATACCTCTCCCTGATGGCATTA 2460

QY 2461 ATGCAGAGTCAGATATCTTCAGGGTTGCTATTGCTGGGGCCCCAGTCACTCTGTGGATC 2520
DB |||||
2461 ATGCAGAGTCAGATATCTTCAGGGTTGCTATTGCTGGGGCCCCAGTCACTCTGTGGATC 2520
QY 2521 TTCTATGATACAGGATACAGGACGTTATATGGTCAACCTTGACACAGATCAACAGGC 2580
DB |||||
2521 TTCTATGATACAGGATACAGGACGTTATATGGTCAACCTTGACACAGATCAACAGGC 2580
QY 2581 TATTACTTATAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTGAAACCAATCGTTTA 2640
DB |||||
2581 TATTACTTATAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTGAAACCAATCGTTTA 2640
QY 2641 CTGCTCTTACATGTTCTCTGGATGAGATGTCATATTTTGCATACACAGATATATCTG 2700
DB |||||
2641 CTGCTCTTACATGTTCTCTGGATGAGATGTCATATTTTGCATACACAGATATATCTG 2700
QY 2701 AGTTTTTATAGGCTGGAAGCCCATATGATTTTACAGATCTATCTCAGGAGACAC 2760
DB |||||
2701 AGTTTTTATAGGCTGGAAGCCCATATGATTTTACAGATCTATCTCAGGAGACAC 2760
QY 2761 AGCATAGAGTTCTGGAATCGGAGAACATATGAACTGCATCTTTTGCATACCTTCAA 2820
DB |||||
2761 AGCATAGAGTTCTGGAATCGGAGAACATATGAACTGCATCTTTTGCATACCTTCAA 2820
QY 2821 GAAACCTTGGATCAGCTATTGCTCTTAAAGTATATATATTTTGCATACCTTCAA 2880
DB |||||
2821 GAAACCTTGGATCAGCTATTGCTCTTAAAGTATATATATTTTGCATACCTTCAA 2880
QY 2881 TCTCTGGTATACCTGCTATTTAAACCAATGAGGAGTTTAAATCAACAGAAAACACAGA 2940
DB |||||
2881 TCTCTGGTATACCTGCTATTTAAACCAATGAGGAGTTTAAATCAACAGAAAACACAGA 2940
QY 2941 ATTGATCATCATCTTTGATACCTGCTGATCACTACTCTCTGAAATTAATGTGGTG 3000
DB |||||
2941 ATTGATCATCATCTTTGATACCTGCTGATCACTACTCTCTGAAATTAATGTGGTG 3000
QY 3001 CATTGACAGGCTTACGGTTTGTGTAGTAACTAAATACCTTAAACCCACATGCTCAAAA 3060
DB |||||
3001 CCATGACAGGCTTACGGTTTGTGTAGTAACTAAATACCTTAAACCCACATGCTCAAAA 3060
QY 3061 TCAATGATACATATTCCTGAGAGACCCAGCAATACATTAAGAAATTAATAAAAAAAA 3120
DB |||||
3061 TCAATGATACATATTCCTGAGAGACCCAGCAATACATTAAGAAATTAATAAAAAAAA 3120

RESULT 2

US-09-976-674-12
; Sequence 12, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976, 674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 4829
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-976-674-12

Query Match 92.6%; Score 2889; DB 3; Length 4829;
Best Local Similarity 95.9%; Pred. No. 0;
Matches 3119; Conservative 0; Mismatches 1; Indels 2; Gaps 1;
QY 1 AAGTGCTAAAGCCTCCGAGGCCAAGGCGCTGCTACTGCGCGCGCTTCTTAGTGGCG 60

DB |||||
1 AAGTGCTAAAGCCTCCGAGGCCAAGGCGCTGCTACTGCGCGCGCTTCTTAGTGGCG 60
QY |||||
61 CGTTCCGCGCTGGGTTGTACCGCGCGCCGCGGAGGAGCACTGCAACACGAGACCG 120
DB |||||
61 CGTTCCGCGCTGGGTTGTACCGCGCGCCGCGGAGGAGCACTGCAACACGAGACCG 120
QY |||||
121 GAGTGGAGCGCGCAGCATGAAGCGCGCAGCGCGCTCCATAGCGCAGCTGCGGACCG 180
DB |||||
121 GAGTGGAGCGCGCAGCATGAAGCGCGCAGCGCGCTCCATAGCGCAGCTGCGGACCG 180
QY |||||
181 TCCGGCGCGCGCGGAGGAAAATGCAACATGCGCAGCAGCAATGGAACACAGAACAG 240
DB |||||
181 TCCGGCGCGCGCGGAGGAAAATGCAACATGCGCAGCAGCAATGGAACACAGAACAG 240
QY |||||
241 CTGGGTGTTGAGATATTTGAAAATGCGCGCACTGTGAGGAGAAATTTGAATACAGGATCGG 300
DB |||||
241 CTGGGTGTTGAGATATTTGAAAATGCGCGCACTGTGAGGAGAAATTTGAATACAGGATCGG 300
QY |||||
301 CCTAAATTTGAGCCCTTTTATGTTGAGCGGTATTTCTGGAGTCAGCTTAAAAAGCTGCTT 360
DB |||||
301 CCTAAATTTGAGCCCTTTTATGTTGAGCGGTATTTCTGGAGTCAGCTTAAAAAGCTGCTT 360
QY |||||
361 GCCGATACAGAAAATATCATGCTCATGATGCTTAAGGCCACCAATGATTTTCAATGTTT 420
DB |||||
361 GCCGATACAGAAAATATCATGCTCATGATGCTTAAGGCCACCAATGATTTTCAATGTTT 420
QY |||||
421 GTGAAGAGGAATGATCCAGATGGAACCTTATTCAGACAGAAATCTTATACCTTGCATGCT 480
DB |||||
421 GTGAAGAGGAATGATCCAGATGGAACCTTATTCAGACAGAAATCTTATACCTTGCATGCT 480
QY |||||
481 GGTGAGAACAGAGAAAATACACTGTTTATTTCTGAAATTTCCCAAAATCTATCAATAGACA 540
DB |||||
481 GGTGAGAACAGAGAAAATACACTGTTTATTTCTGAAATTTCCCAAAATCTATCAATAGACA 540
QY |||||
541 GCAGTCTTAATGCTCTCTTGGAAAGCTCTTTTGGATCTTTTTCAGGCAACACTTGAAGTAT 600
DB |||||
541 GCAGTCTTAATGCTCTCTTGGAAAGCTCTTTTGGATCTTTTTCAGGCAACACTTGAAGTAT 600
QY |||||
601 GGAATGATTTCTCAGAGAAAGAACTATTAAGAGAAAGAAACGCATTTGGAACAGTCGGA 660
DB |||||
601 GGAATGATTTCTCAGAGAAAGAACTATTAAGAGAAAGAAACGCATTTGGAACAGTCGGA 660
QY |||||
661 ATTGCTTCTTACAGTATATCACCAGGAAGTGGAAACATTTCTGTTTCAAGCCGCTAGTGA 720
DB |||||
661 ATTGCTTCTTACAGTATATCACCAGGAAGTGGAAACATTTCTGTTTCAAGCCGCTAGTGA 720
QY |||||
721 ATTTATACAGTAAAGATGGAGGCCAAGAGATTTTACGCAACAACTTTTAAAGGCCCAAT 780
DB |||||
721 ATTTATACAGTAAAGATGGAGGCCAAGAGATTTTACGCAACAACTTTTAAAGGCCCAAT 780
QY |||||
781 CTAGTGGAACACTAGTGTCTCCCAACATACGATGATCCAAATTTATGCCCCCTGCTATCCA 840
DB |||||
781 CTAGTGGAACACTAGTGTCTCCCAACATACGATGATCCAAATTTATGCCCCCTGCTATCCA 840
QY |||||
841 GACTGGATGCTTTTATACATAGCAACGATATTTGGATATCTAACAATGTAACAGAGAA 900
DB |||||
841 GACTGGATGCTTTTATACATAGCAACGATATTTGGATATCTAACAATGTAACAGAGAA 900
QY |||||
901 GAAAGGAGACTCACTTATGTGCAATGAGCTAGCCAACTGGAAGAGATGCCAGATCA 960
DB |||||
901 GAAAGGAGACTCACTTATGTGCAATGAGCTAGCCAACTGGAAGAGATGCCAGATCA 960
QY |||||
961 GCTGGAGTCGCTACCTTTGTTCTCCAGAGAGAAATTTGATAGATATTTCTGGCTATTGGTGG 1020
DB |||||
961 GCTGGAGTCGCTACCTTTGTTCTCCAGAGAGAAATTTGATAGATATTTCTGGCTATTGGTGG 1020
QY |||||
1021 TGTCCAAAGCTGAAACAACTCCCGAGTGGTGGTAAATTTCTTAGAATTTCTATATGAAGA 1080
DB |||||
1021 TGTCCAAAGCTGAAACAACTCCCGAGTGGTGGTAAATTTCTTAGAATTTCTATATGAAGA 1080
QY |||||
1081 AATGATGAATCTGAGGTGGAAATTTATTTTACATCCCTATGTTGGAAACAAGAGG 1140

Db 1081 AATGATGAACTCTGAGGTGGAAATTAATTCATGTTTACATCCCTATCTGTTGGAACAAGGAGG 1140
Qy 1141 GCAGATCAATTCGGTTATCTTAACACAGGTACAGCAAAATCCCTAAAGTCACATTTTAAAGATG 1200
Db 1141 GCAGATCAATTCGGTTATCTTAACACAGGTACAGCAAAATCCCTAAAGTCACATTTTAAAGATG 1200
Qy 1201 TCAGAAATAATGATGTGATGTCGAAGGAAGGATCATAGATGTCATAGATAAGAACTAATTT 1260
Db 1201 TCAGAAATAATGATGTGATGTCGAAGGAAGGATCATAGATGTCATAGATAAGAACTAATTT 1260
Qy 1261 CAACCTTTTGAGATCTTATTTTGAAGGAGTTGAATATATTGCCAGAGCTGGATGGACCTCT 1320
Db 1261 CAAGCTTTTGAGATCTTATTTGAAGGAGTTGAATATATTGCCAGAGCTGGATGGACCTCT 1320
Qy 1321 GAGGAAAAATATGCTTGGTCCATCCTACTAGATCGCTCCAGACCTCGCTACAGATAGTG 1380
Db 1321 GAGGAAAAATATGCTTGGTCCATCCTACTAGATCGCTCCAGACCTCGCTACAGATAGTG 1380
Qy 1381 TTGATCTCACTGAAATTAATTTATCCAGTAGAAGATGATGTTATGGAAGGCAGAGACTC 1440
Db 1381 TTGATCTCACTGAAATTAATTTATCCAGTAGAAGATGATGTTATGGAAGGCAGAGACTC 1440
Qy 1441 ATTGAGTCAGTGCCTGATCTGTGACGCCACTAAATTTATCTATGAAGAAACAACAGACATC 1500
Db 1441 ATTGAGTCAGTGCCTGATCTGTGACGCCACTAAATTTATCTATGAAGAAACAACAGACATC 1500
Qy 1501 TGGATAAATATCCATGACATCTTTTCATGTTTCCCAAGTCACGAGAGGAAATTTGAG 1560
Db 1501 TGGATAAATATCCATGACATCTTTTCATGTTTCCCAAGTCACGAGAGGAAATTTGAG 1560
Qy 1561 TTTATTTTGGCTCTGAAATCAAAAGAGTTTCCGTCATTTATACAAAATTTACATCTATT 1620
Db 1561 TTTATTTTGGCTCTGAAATCAAAAGAGTTTCCGTCATTTATACAAAATTTACATCTATT 1620
Qy 1621 TTAAGGAAGCAAAATTAACGATCCAGTGTGGCTGCTGCTCCAAAGTGATTTCAAG 1680
Db 1621 TTAAGGAAGCAAAATTAACGATCCAGTGTGGCTGCTGCTCCAAAGTGATTTCAAG 1680
Qy 1681 TGTCCTATCAAAGAGGAGATAGCAATTTACAGTGTGTAATGGGAAGTTCTTTGGCCGGCAT 1740
Db 1681 TGTCCTATCAAAGAGGAGATAGCAATTTACAGTGTGTAATGGGAAGTTCTTTGGCCGGCAT 1740
Qy 1741 GGATCTAATATCCAAGTTGATGAAGTCAGAAAGCTGGTATATTTTGAAGGCCACCAAGAC 1800
Db 1741 GGATCTAATATCCAAGTTGATGAAGTCAGAAAGCTGGTATATTTTGAAGGCCACCAAGAC 1800
Qy 1801 TCCCTTTTAGAGCATCACCTGTACGTAGTCAGTTACGTAAATCCTGGAGAGGTGACAAAG 1860
Db 1801 TCCCTTTTAGAGCATCACCTGTACGTAGTCAGTTACGTAAATCCTGGAGAGGTGACAAAG 1860
Qy 1861 CTGACTGACCGTGGCTACTCACATTTCTGCTGCATCAGTCAGCAGCTGTGACTTCTTTATA 1920
Db 1861 CTGACTGACCGTGGCTACTCACATTTCTGCTGCATCAGTCAGCAGCTGTGACTTCTTTATA 1920
Qy 1921 AGTAAGTATAGTAACCAAGAGAAATCCACACTGTGTGCTCCCTTTACAAGCTATCAAGTCT 1980
Db 1921 AGTAAGTATAGTAACCAAGAGAAATCCACACTGTGTGCTCCCTTTACAAGCTATCAAGTCT 1980
Qy 1981 GAAGATGACCCAACTTGCRAAACAAAGGAATTTTGGGCCACATTTTGGATTCAGCAGGT 2040
Db 1981 GAAGATGACCCAACTTGCRAAACAAAGGAATTTTGGGCCACATTTTGGATTCAGCAGGT 2040
Qy 2041 CCTCTTCTGACTATCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGATTTTACA 2100
Db 2041 CCTCTTCTGACTATCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGATTTTACA 2100
Qy 2101 TTGATGGGATGCTCTACAAGCCTCATGATCTACAGCTCGGAAGAAATATCCTACTGTG 2160
Db 2101 TTGATGGGATGCTCTACAAGCCTCATGATCTACAGCTCGGAAGAAATATCCTACTGTG 2160
Qy 2161 CTGTTTCATATATGGTGG--TCCTCAGGTGAGTTGGTGAATATCGGTTTAAAGGAGTCA 2218
Db 2161 CTGTTTCATATATGGTGGTCTCCTCAGGTGAGTTGGTGAATATCGGTTTAAAGGAGTCA 2220

Qy 2219 AGTATTTCCGCTTGAATACCCCTAGCCTCTCTAGGTTTATGTTGTTAGTGATAGACAACA 2278
Db 2221 AGTATTTCCGCTTGAATACCCCTAGCCTCTCTAGGTTTATGTTGTTAGTGATAGACAACA 2280
Qy 2279 GGGGATCCTGTACCGAGGGCTTTAAATTTGAAGGCCCTTTTAAATATATAAATGGGTCAAA 2338
Db 2281 GGGGATCCTGTACCGAGGGCTTTAAATTTGAAGGCCCTTTTAAATATATAAATGGGTCAAA 2340
Qy 2339 TAGAAATTTGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTCAATG 2398
Db 2341 TAGAAATTTGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTCAATG 2400
Qy 2399 ACTTAGATCTGTGTGGGCATCCACGGCTGTCTCTATGGAGATACCTCTCCCTGATGGCAT 2458
Db 2401 ACTTAGATCTGTGTGGGCATCCACGGCTGTCTCTATGGAGATACCTCTCCCTGATGGCAT 2460
Qy 2459 TAATGTCAGAGGTACAGATATCTTCAGGGTGTCTATTGCTGGGGCCCCAGTCACTCTGTGGA 2518
Db 2461 TAATGTCAGAGGTACAGATATCTTCAGGGTGTCTATTGCTGGGGCCCCAGTCACTCTGTGGA 2520
Qy 2519 TCTTCTATGATACAGGATACACGGAACGTTTATATGGGTCACTCTGACAGAAATGAACAGG 2578
Db 2521 TCTTCTATGATACAGGATACACGGAACGTTTATATGGGTCACTCTGACAGAAATGAACAGG 2580
Qy 2579 GCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAGTTCCCTCTGAAACCAATCGTT 2638
Db 2581 GCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAGTTCCCTCTGAAACCAATCGTT 2640
Qy 2639 TACTGCTCTTACATGTTTCTCGGATGAGAAATGTCCTTTTGCACATACCAAGTATATTAC 2698
Db 2641 TACTGCTCTTACATGTTTCTCGGATGAGAAATGTCCTTTTGCACATACCAAGTATATTAC 2700
Qy 2699 TGAGTTTTTTAGTGAGGGCTGGAAAGCCATATGATTTTACAGATCTATCTCTCAGGAGAGAC 2758
Db 2701 TGAGTTTTTTAGTGAGGGCTGGAAAGCCATATGATTTTACAGATCTATCTCTCAGGAGAGAC 2760
Qy 2759 ACAGATAAGAGTTTCTGAAATCGGGAGAACATATGAACCTGTCATCTTTTGCACTACCTTC 2818
Db 2761 ACAGATAAGAGTTTCTGAAATCGGGAGAACATATGAACCTGTCATCTTTTGCACTACCTTC 2820
Qy 2819 AAGAAAACCTTGGATCACGTTATTTGCTCTAAAAGTGATATAATTTTGAOCTGTGTAGA 2878
Db 2821 AAGAAAACCTTGGATCACGTTATTTGCTCTAAAAGTGATATAATTTTGAOCTGTGTAGA 2880
Qy 2879 ACTCTCTGTATACACTGGCTATTTAAACCAAAATGAGGAGGTTTAAATCAACAGAAACACA 2938
Db 2881 ACTCTCTGTATACACTGGCTATTTAAACCAAAATGAGGAGGTTTAAATCAACAGAAACACA 2940
Qy 2939 GAATTGATCATCACATTTTGTATACCTGCCATGTAACATCTACTCTCTGAAAATFAAATGTGG 2998
Db 2941 GAATTGATCATCACATTTTGTATACCTGCCATGTAACATCTACTCTCTGAAAATFAAATGTGG 3000
Qy 2999 TGCCATGACGGGTCTACGGTTTGTGTAGTAAATCTTAATACCTTAACCCCATGCTCAA 3058
Db 3001 TGCCATGACGGGTCTACGGTTTGTGTAGTAAATCTTAATACCTTAACCCCATGCTCAA 3060
Qy 3059 AATCAAAATGATACATATTTCTGTAGAGACCCAGCAATACCATAGAAATTAACAAAAAAA 3118
Db 3061 AATCAAAATGATACATATTTCTGTAGAGACCCAGCAATACCATAGAAATTAACAAAAAAA 3120
Qy 3119 AA 3120
Db 3121 AA 3122

RESULT 3

US-09-976-674-2
; Sequence 2, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen

; APPLICANT: Riviere, Pierre									
; APPLICANT: Junien, Jean-Louis									
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV									
; FILE OF INVENTION: 70669									
; CURRENT APPLICATION NUMBER: US/09/976,674									
; CURRENT FILING DATE: 2001-10-12									
; PRIOR APPLICATION NUMBER: US 60/240,117									
; PRIOR FILING DATE: 2000-10-12									
; NUMBER OF SEQ ID NOS: 61									
; SOFTWARE: PatentIn version 3.1									
; SEQ ID NO 2									
; LENGTH: 2671									
; TYPE: DNA									
; ORGANISM: Homo sapiens									
US-09-976-674-2									
Query Match 80.0%; Score 2495; DB 3; Length 2671;									
Best Local Similarity 99.9%; Pred No. 0;									
Matches 2645; Conservative 0; Mismatches 3; Indels 0; Gaps 0;									
Qy	213	CATGGCAGCAGCAATGGAAACAGAACAGCTGGGTGTGAGATATTTGAAACTCGGACTG	272						
Db	7	CATGGCAGCAGCAATGGAAACAGAACAGCTGGGTGTGAGATATTTGAAACTCGGACTG	66						
Qy	273	TGAGGAGAAATTTGAATCA CAGGATCGGCCTAAATTTGAGGCTTTTATGTTGAGCGGTA	332						
Db	67	TGAGGAGAAATTTGAATCA CAGGATCGGCCTAAATTTGAGGCTTTTATGTTGAGCGGTA	126						
Qy	333	TTCTCGAGTCAGCTTAAAAAGCTGTTGCCGATACAGAAATATCATGCTACATGAT	392						
Db	127	TTCTCGAGTCAGCTTAAAAAGCTGTTGCCGATACAGAAATATCATGCTACATGAT	186						
Qy	393	GGCTAAGGCACCATGATTTTATGTTGTGAAGGAGATCATCCAGATGGACTCATTC	452						
Db	187	GGCTAAGGCACCATGATTTTATGTTGTGAAGGAGATCATCCAGATGGACTCATTC	246						
Qy	453	AGACAGAACTCTATTACTTGGCCATGCTGGTGAGAACAGAGAAATACACTGTTTTATTTC	512						
Db	247	AGACAGAACTCTATTACTTGGCCATGCTGGTGAGAACAGAGAAATACACTGTTTTATTTC	306						
Qy	513	TGAAATTTCCAAATCTATCAATAGACAGCAGCTCTTAATGCTCTCTTTGGAAGCTCTTTT	572						
Db	307	TGAAATTTCCAAATCTATCAATAGACAGCAGCTCTTAATGCTCTCTTTGGAAGCTCTTTT	366						
Qy	573	GGATCTTTTTCAGGCAACACTGGAATATGTAATGATTTCTCGAGAGAGAACTATTAAG	632						
Db	367	GGATCTTTTTCAGGCAACACTGGAATATGTAATGATTTCTCGAGAGAGAACTATTAAG	426						
Qy	633	AGAAAGAAACGGATTTGGAACAGTCGGAATTCCTTACGATTTATCACCAGGAAGTGG	692						
Db	427	AGAAAGAAACGGATTTGGAACAGTCGGAATTCCTTACGATTTATCACCAGGAAGTGG	486						
Qy	693	AACATTTCTGTTTTCAGCCGGTAGTGGAAATTTATCACGTAAGATGGAGGCCACAAAG	752						
Db	487	AACATTTCTGTTTTCAGCCGGTAGTGGAAATTTATCACGTAAGATGGAGGCCACAAAG	546						
Qy	753	ATTTTACGCAACAACTTTTAAGGCCAATCTAGTGGAAATCTAGTTGTCCTCAACATACCGAT	812						
Db	547	ATTTTACGCAACAACTTTTAAGGCCAATCTAGTGGAAATCTAGTTGTCCTCAACATACCGAT	606						
Qy	813	GGATCCAAAATTTATGCCCGCTGATCCAGACTGGATTTGCTTTATACATAGCAACGATAT	872						
Db	607	GGATCCAAAATTTATGCCCGCTGATCCAGACTGGATTTGCTTTATACATAGCAACGATAT	666						
Qy	873	TTGGATATCTAACATCGTAAACAGAGAGAAAGGAGACTCACTTATGTCACAATGAGCT	932						
Db	667	TTGGATATCTAACATCGTAAACAGAGAGAAAGGAGACTCACTTATGTCACAATGAGCT	726						
Qy	933	AGCCAAATCGGAGAGAGATGCCAGATCAGCTGGATTCGCTACCTTTGTTCTCCAGAGA	992						
Db	727	AGCCAAATCGGAGAGAGATGCCAGATCAGCTGGATTCGCTACCTTTGTTCTCCAGAGA	786						
Qy	993	ATTTGATAGATATTTCTGGCTATTGGTGGTGTCCAAAGCTGAAACAACTCCCGAGTGGTGG	1052						

Db	787	ATTGTAGATATTTCTGGCTATTGGTGGTGTCCAAAGCTGAAACAACTCCCGTGGTGG	846						
Qy	1053	TAAATTTCTTAGAATTTCTATATGAAGAAATGAATCTCAGGTGGAATTTATTCATGT	1112						
Db	847	TAAATTTCTTAGAATTTCTATATGAAGAAATGAATCTCAGGTGGAATTTATTCATGT	906						
Qy	1113	TACATCCCTATGTTGGAACCAAGGAGGCGAGATTCATTCGGTTATCTTAAACAGGTAC	1172						
Db	907	TACATCCCTATGTTGGAACCAAGGAGGCGAGATTCATTCGGTTATCTTAAACAGGTAC	966						
Qy	1173	AGCAAAATCTTAAAGTCACTTTTAAAGATGTCAGAAATAATGATGATGCTGAAGAAAGAT	1232						
Db	967	AGCAAAATCTTAAAGTCACTTTTAAAGATGTCAGAAATAATGATGATGCTGAAGAAAGAT	1026						
Qy	1233	CATAGATGTCATAGATAAGGAACCTAATTCACCTTTTGAGATTTCTATTTGAAGAGTTGA	1292						
Db	1027	CATAGATGTCATAGATAAGGAACCTAATTCACCTTTTGAGATTTCTATTTGAAGAGTTGA	1086						
Qy	1293	ATATATTGCCAGAGCTGGATGGACTCCTGAGGGAATAATGCTTGGTCCATCTACTAGA	1352						
Db	1087	ATATATTGCCAGAGCTGGATGGACTCCTGAGGGAATAATGCTTGGTCCATCTACTAGA	1146						
Qy	1353	TCGCTCCAGACTCGCTTACAGATAGTGTGATCTCACTGAAATTTATTTATCCAGTAGA	1412						
Db	1147	TCGCTCCAGACTCGCTTACAGATAGTGTGATCTCACTGAAATTTATTTATCCAGTAGA	1206						
Qy	1413	AGATGATGTTTATGGAAGGCGAGACTCAATTCAGTCAGTGGCTGATTTCTGTGAGCCACT	1472						
Db	1207	AGATGATGTTTATGGAAGGCGAGACTCAATTCAGTCAGTGGCTGATTTCTGTGAGCCACT	1266						
Qy	1473	AAATATCTATGGAAGAAACACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTT	1532						
Db	1267	AAATATCTATGGAAGAAACACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTT	1326						
Qy	1533	TCGCCAAGTCAAGAGAGAAATTTGAGTTTATTTTGGCTCTGAATGCCAAAACAGGTTT	1592						
Db	1327	TCGCCAAGTCAAGAGAGAAATTTGAGTTTATTTTGGCTCTGAATGCCAAAACAGGTTT	1386						
Qy	1593	CCGTCATTTTACAAAATTTACATCTATTTTAAAGAAACCAATATAACCATCCAGTGG	1652						
Db	1387	CCGTCATTTTACAAAATTTACATCTATTTTAAAGAAACCAATATAACCATCCAGTGG	1446						
Qy	1653	TGGCTCCCTGCTCCAAAGTGTTCCTCAAGAGGAGATAGCAATTTACCAG	1712						
Db	1447	TGGCTCCCTGCTCCAAAGTGTTCCTCAAGAGGAGATAGCAATTTACCAG	1506						
Qy	1713	TGTTGAATGGGAAGTTCTTTGGCCGCGATGATTAATATCCAAGTTGATGAAGTCAGAAG	1772						
Db	1507	TGTTGAATGGGAAGTTCTTTGGCCGCGATGATTAATATCCAAGTTGATGAAGTCAGAAG	1566						
Qy	1773	GCTGGTATATTTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTACGTAGTCAG	1832						
Db	1567	GCTGGTATATTTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTACGTAGTCAG	1626						
Qy	1833	TTACGTAATTCCTGGAGAGGTGACAAAGGTGACTGACCGTGGCTACTCACATTTCTGCTG	1892						
Db	1627	TTACGTAATTCCTGGAGAGGTGACAAAGGTGACTGACCGTGGCTACTCACATTTCTGCTG	1686						
Qy	1893	CATCAGTCAGACTGTGACTTCTTTTATAGTAAGTATAGTACACAGAGAATCCACACTG	1952						
Db	1687	CATCAGTCAGACTGTGACTTCTTTTATAGTAAGTATAGTACACAGAGAATCCACACTG	1746						
Qy	1953	TGTTGCTCTTTTACAAAGCTATCAAGTCTCTGAGATGACCAACTTGCAGAAAACAAAGGAAT	2012						
Db	1747	TGTTGCTCTTTTACAAAGCTATCAAGTCTCTGAGATGACCAACTTGCAGAAAACAAAGGAAT	1806						
Qy	2013	TTGGGCCACCACTTTTGGATTCAGAGGTCTCTCTCTGCTACTATCTCTCCAGAAATTTT	2072						
Db	1807	TTGGGCCACCACTTTTGGATTCAGAGGTCTCTCTCTGCTACTATCTCTCCAGAAATTTT	1866						
Qy	2073	CTCTTTTGAAGTACTACTGGAATTTACATTTGATGGATGCTCTTACAAGCCCTCATGATCT	2132						

Db 1867 CTCTTTTGAAGTACTACTGGATTACATTTGATGGATGCTCTACAGGCTCATGATCT 1926
QY 2133 ACAGCCTGGAAAGAAATATCCTACTGTGCTGTTTCATATATGTTGGTCTCTAGTGCGAGTT 2192
Db 1927 ACAGCCTGGAAAGAAATATCCTACTGTGCTGTTTCATATATGTTGGTCTCTAGTGCGAGTT 1986
QY 2193 GGTGAATAATCGGCTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCCCTAGGCTCTCAGG 2252
Db 1987 GGTGAATAATCGATTTAAAGGAGTCAAGTATTTCCGCTTGAATACCCCTAGGCTCTCAGG 2046
QY 2253 TTATGTGGTGTAGTATGATAGACAAAGGGGATCCTGTGCCAGGGGCTTTAAATTTGAAG 2312
Db 2047 TTATGTGGTGTAGTATGATAGACAAAGGGGATCCTGTGCCAGGGGCTTTAAATTTGAAG 2106
QY 2313 CGCCTTTAAATATAAATGGGTCAATATAGAAATGACGATCAGGTGGAAGGACTCCAAATA 2372
Db 2107 CGCCTTTAAATATAAATGGGTCAATATAGAAATGACGATCAGGTGGAAGGACTCCAAATA 2166
QY 2373 TCTAGCTTCTCGATATGATTTCAITGACTTAGATCGTGTGGGCATCCACGGCTGGTCCCTA 2432
Db 2167 TCTAGCTTCTCGATATGATTTCAITGACTTAGATCGTGTGGGCATCCACGGCTGGTCCCTA 2226
QY 2433 TGGAGGATACCTCTCCCTGATGGCATTAATGACAGAGTTCAGATATCTTTCAGGGTTGCTAT 2492
Db 2227 TGGAGGATACCTCTCCCTGATGGCATTAATGACAGAGTTCAGATATCTTTCAGGGTTGCTAT 2286
QY 2493 TGTGTGGGGCCCCAGTCACCTCTGTGGATCTTCTATGATACAGGATACACGGAAGCTTATAT 2552
Db 2287 TGTGTGGGGCCCCAGTCACCTCTGTGGATCTTCTATGATACAGGATACACGGAAGCTTATAT 2346
QY 2553 GGGTCACCTGACCCAGATGACAGGCTATTAAGTATGATACAGGATACACGGAAGCTTATAT 2612
Db 2347 GGGTCACCTGACCCAGATGACAGGCTATTAAGTATGATACAGGATACACGGAAGCTTATAT 2406
QY 2613 AAGTTTCCCTCTGACCCAAATCGTTTACTGCTCTTACATGTTTCTTGATGAGAAATGT 2672
Db 2407 AAGTTTCCCTCTGACCCAAATCGTTTACTGCTCTTACATGTTTCTTGATGAGAAATGT 2466
QY 2673 CCATTTTGACATACAGTATATTAAGTATGATGAGGCTGGAAGCCATATGA 2732
Db 2467 CCATTTTGACATACAGTATATTAAGTATGATGAGGCTGGAAGCCATATGA 2526
QY 2733 TTATACAGATCTATCTTCAGGAGAGACAGCATTAAGAGTTTCTTGAATTCGGGAGAACATTA 2792
Db 2527 TTATACAGATCTATCTTCAGGAGAGACAGCATTAAGAGTTTCTTGAATTCGGGAGAACATTA 2586
QY 2793 TGAATCTGATCTTTTGGCACTACCTTCAAGAAAACCTTTGGATCAGTATTCGCTCTAAA 2852
Db 2587 TGAATCTGATCTTTTGGCACTACCTTCAAGAAAACCTTTGGATCAGTATTCGCTCTAAA 2646
QY 2853 AGTGATAT 2860
Db 2647 AGTGATAT 2654

RESULT 4
US-09-976-594-1103
; Sequence 1103, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
; APPLICANT: Fumess, Michael
; APPLICANT: Buchinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCE: PA-0041 US
; CURRENT FILING DATE: 2001-10-12
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 1103
; LENGTH: 2797
; TYPE: DNA

; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. 6673549 977951.1
US-09-976-594-1103

Query Match 74.2%; Score 2315; DB 3; Length 2797;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2785; Conservative 0; Mismatches 1; Indels 5; Gaps 2;

QY 335 CCTGGAGTCAGCTTAAAAAGCTGCTTGGCGATACAGAAAAATATCATGGCTACATGATGG 394
Db 1 CCTGGAGTCAGCTTAAAAAGCTGCTTGGCGATACAGAAAAATATCATGGCTACATGATGG 60
QY 395 CTAAGGCACCATGATGATTTTCATGTTTGTGAAGAGAAATGATCCAGATGGACCTCATTCAG 454
Db 61 CTAAGGCACCATGATGATTTTCATGTTTGTGAAGAGAAATGATCCAGATGGACCTCATTCAG 120
QY 455 ACAGAAATCTATTACCTTGGCCATGCTGTGGTGAACAGAGAAATACACTGTTTATTCGTG 514
Db 121 ACAGAAATCTATTACCTTGGCCATGCTGTGGTGAACAGAGAAATACACTGTTTATTCGTG 180
QY 515 AAATTTCCCAAAATCAATAGACGACGAGTCTTAATGCTCTCTTGGAGCCTCTTTGG 574
Db 181 AAATTTCCCAAAATCAATAGACGACGAGTCTTAATGCTCTCTTGGAGCCTCTTTGG 240
QY 575 ATCTTTTTCAGGCAACACATGGACTATGGAATGTTATTCGAGAAAGAACTATTAAGAG 634
Db 241 ATCTTTTTCAGGCAACACATGGACTATGGAATGTTATTCGAGAAAGAACTATTAAGAG 300
QY 635 AAAGAAAACGCAATGGAACAGTCGGAATGCTTCTTACGATTTATCACCAAGGAAGTGAA 694
Db 301 AAAGAAAACGCAATGGAACAGTCGGAATGCTTCTTACGATTTATCACCAAGGAAGTGAA 360
QY 695 CATTTCTGTTTCAAGCCGGTAGTGGAAATTTATACGTAAGAGATGGAGGCCCAAGGAT 754
Db 361 CATTTCTGTTTCAAGCCGGTAGTGGAAATTTATACGTAAGAGATGGAGGCCCAAGGAT 420
QY 755 TTACGCAACCAACTTTTAAAGGCCCAATCTAGTGGGAAACTAGTTTGTCCCAACATACGATAT 874
Db 421 TTACGCAACCAACTTTTAAAGGCCCAATCTAGTGGGAAACTAGTTTGTCCCAACATACGATAT 480
QY 875 ATCCAAAAATATGCCCCGCTGATCCAGACTGGATTTGCTTTTATACATAGCAACGATATTT 874
Db 481 ATCCAAAAATATGCCCCGCTGATCCAGACTGGATTTGCTTTTATACATAGCAACGATATTT 540
QY 875 GGATATCTAAACATCGTAACCCAGAGAAAGAGACTCACTTATGTGCACAATGAGCTAG 934
Db 541 GGATATCTAAACATCGTAACCCAGAGAAAGAGACTCACTTATGTGCACAATGAGCTAG 600
QY 935 CCACATATGGAAGAAAGATGCCAGATCAGCTGGAGTCGCTAGCTTTGTTCTCCAGAGAAAT 994
Db 601 CCACATATGGAAGAAAGATGCCAGATCAGCTGGAGTCGCTAGCTTTGTTCTCCAGAGAAAT 660
QY 995 TTGATAGATATTTCTGGCTATTTGGTGGTCCAAAGCTGAAACAACTCCAGTGGTGGTA 1054
Db 661 TTGATAGATATTTCTGGCTATTTGGTGGTCCAAAGCTGAAACAACTCCAGTGGTGGTA 720
QY 1055 AAATTTCTAGAAATCTATATGAAGAAATGATGAATCTGAGGTGGAAATTTATTCATGTTA 1114
Db 721 AAATTTCTAGAAATCTATATGAAGAAATGATGAATCTGAGGTGGAAATTTATTCATGTTA 780
QY 1115 CATCCCTATGTTGGAACAAAGG-AGGCGAGATTTCATTCCGTTATCTCTAAAAACAGGTACA 1173
Db 781 CATCCCTATGTTGGAACAAAGGCGAGGCGAGATTTCATTCCGTTATCTCTAAAAACAGGTACA 840
QY 1174 GCAAAATCCTAAAGTCACCTTTTAAAGATGTCAAGAAATATGATGCTGAAGGAAGGATC 1233
Db 841 GCAAAATCCTAAAGTCACCTTTTAAAGATGTCAAGAAATATGATGCTGAAGGAAGGATC 900
QY 1234 ATAGATGTCAATAGTAAGGAACTAAATCAACCTTTTGGAGATTTCTATTTGAAGGAGTTGAA 1293
Db 901 ATAGATGTCAATAGTAAGGAACTAAATCAACCTTTTGGAGATTTCTATTTGAAGGAGTTGAA 960

1294 TATATTGCCAGAGCTGGATGACTCTCGAGGAAAAATA-----TGCTTGGTCCATCTCTACT 1349
Db TATATTGCCAGAGCTGGATGACTCTCGAGGAAAAATAGTGCTTGGTCCATCTCTACT 1020
1350 AGATCGCTCCCAAGACTCGCCTACAGATAGTGTGATCTCACTGGAATATTATATCCAGT 1409
Db AGATCGCTCCCAAGACTCGCCTACAGATAGTGTGATCTCACTGGAATATTATATCCAGT 1080
1410 AGAAGATGATGTATTGGAAGGCAGAGACTCAATTGAGTCACTGCTGATCTGTGACGCC 1469
Db AGAAGATGATGTATTGGAAGGCAGAGACTCAATTGAGTCACTGCTGATCTGTGACGCC 1140
1470 ACTAATATTCTATGAAGAAAAACAAGACATCTGATATAATATCCATGACATCTTTTCATGT 1529
Db ACTAATATTCTATGAAGAAAAACAAGACATCTGATATAATATCCATGACATCTTTTCATGT 1200
1530 TTTTCCCAAGACTCAAGAGGAAATTTGAGTTATTTTTTGGCTCTGAAATGCAAAACAGG 1589
Db TTTTCCCAAGACTCAAGAGGAAATTTGAGTTATTTTTTGGCTCTGAAATGCAAAACAGG 1260
1590 TTTCGGTCAATTTATCAAAATTTACATCTATTTTAAAGGAAGCAATATAAAGCATCCAG 1649
Db TTTCGGTCAATTTATCAAAATTTACATCTATTTTAAAGGAAGCAATATAAAGCATCCAG 1320
1650 TGGTGGGCTGCTGCTCCAAAGTGATTTCAAGTGTCTTCAAGAGGAGATAGCAATTTAC 1709
Db TGGTGGGCTGCTGCTCCAAAGTGATTTCAAGTGTCTTCAAGAGGAGATAGCAATTTAC 1380
1710 CAGTGTGTAATGGGAAGTTCTTGGCGGCATGGAATCTAATATCAAGTTGATGAATCAG 1769
Db CAGTGTGTAATGGGAAGTTCTTGGCGGCATGGAATCTAATATCAAGTTGATGAATCAG 1440
1770 AAGGCTGGTATATTTTGAAGGCAAAAGACTCCCTTTAGAGCATCAGCTGACCTAGT 1829
Db AAGGCTGGTATATTTTGAAGGCAAAAGACTCCCTTTAGAGCATCAGCTGACCTAGT 1500
1830 CAGTTACGTAATCTTGGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACAATCTTGG 1889
Db CAGTTACGTAATCTTGGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACAATCTTGG 1560
1890 CTGCATCAGTCAGCACTGTGACTTCTTTATAAGTAAGTATAGTAACCAAGAAATCAACA 1949
Db CTGCATCAGTCAGCACTGTGACTTCTTTATAAGTAAGTATAGTAACCAAGAAATCAACA 1620
1950 CTGTGTGCTCCTTTACAGCTATCAAGTCTTGAAGATGACCCCACTTGCRAAACAAGGA 2009
Db CTGTGTGCTCCTTTTACAGCTATCAAGTCTTGAAGATGACCCCACTTGCRAAACAAGGA 1680
2010 ATTTTGGGCCACCATTTTGGATTCAGAGGTCCTCTTCTGACTATATCTCTCCAGAAAT 2069
Db ATTTTGGGCCACCATTTTGGATTCAGAGGTCCTCTTCTGACTATATCTCTCCAGAAAT 1740
2070 TTTCTCTTTTGAAGTACTACTGGAATTTACATGTPATGGGATGCTCTACAGGCTCATGA 2129
Db TTTCTCTTTTGAAGTACTACTGGAATTTACATGTPATGGGATGCTCTACAGGCTCATGA 1800
2130 TCTACAGCTGGAAGAAATATCTTACTGTGTGCTTTCATATATGCTGTGCTCCTCAGGTGCA 2189
Db TCTACAGCTGGAAGAAATATCTTACTGTGTGCTTTCATATATGCTGTGCTCCTCAGGTGCA 1860
2190 GTTGGTGAATATCGGTTTAAAGGAGTCAAGTATTTTCCGCTTGAATACCCCTAGCCTCTCT 2249
Db GTTGGTGAATATCGGTTTAAAGGAGTCAAGTATTTTCCGCTTGAATACCCCTAGCCTCTCT 1920
2250 AGGTTATGTGGTTGTAGTATAGACAAACAGGGGATCTGTCCAGGAGGCTTAAATTTGA 2309
Db AGGTTATGTGGTTGTAGTATAGACAAACAGGGGATCTGTCCAGGAGGCTTAAATTTGA 1980
2310 AGGCGCTTTAAATATAAATGGGTCAAAATAGAAATTCACCATCAGGTGGAGGACTCCA 2369
Db AGGCGCTTTAAATATAAATGGGTCAAAATAGAAATTCACCATCAGGTGGAGGACTCCA 2040

2370 ATATCTAGCTTCTCGATATGATTTCAATTTAGACTTAGATCGTGTGGGCATCCACGGCTGGTC 2429
Db ATATCTAGCTTCTCGATATGATTTCAATTTAGACTTAGATCGTGTGGGCATCCACGGCTGGTC 2100
2430 CTATGAGGATACCTCTCCCTGATGGCATTAATACAGAGGTACAGATATCTTCAAGGTTGC 2489
Db CTATGAGGATACCTCTCCCTGATGGCATTAATACAGAGGTACAGATATCTTCAAGGTTGC 2160
2490 TATTGCTGGGGCCCAAGTCACTCTGTGATCTTCTATGATACAGGATACACGGAACGTTA 2549
Db TATTGCTGGGGCCCAAGTCACTCTGTGATCTTCTATGATACAGGATACACGGAACGTTA 2220
2550 TATGGGTCAACCTGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGC 2609
Db TATGGGTCAACCTGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGC 2280
2610 AGAAAAGTTCCCTCTGAAACCAAAATCGTTTACTGCTTTACATGCTTTCTCTGATGAGAA 2669
Db AGAAAAGTTCCCTCTGAAACCAAAATCGTTTACTGCTTTACATGCTTTCTCTGATGAGAA 2340
2670 TGTCCATTTTGCACATACCAAGTATATTACTGAGTTTTTTAGTGAGGCTGGAAAGCCATA 2729
Db TGTCCATTTTGCACATACCAAGTATATTACTGAGTTTTTTAGTGAGGCTGGAAAGCCATA 2400
2730 TGATTTACAGATCTATCTCAGGAGAGACACAGCATTAAGAGTTCTGAAATCGGAGAAC 2789
Db TGATTTACAGATCTATCTCAGGAGAGACACAGCATTAAGAGTTCTGAAATCGGAGAAC 2460
2790 TTAGTAACCTGCACTTTTGGCACTTCTCAAGAAAACCTTGGATCAGTATTTGCTGCTCT 2849
Db TTAGTAACCTGCACTTTTGGCACTTCTTCAAGAAAACCTTGGATCAGTATTTGCTGCTCT 2520
2850 AAAAGTGATATAAATTTGACCTGTGTAGAACTCTCTGGTATACACTGGCTATTTAAACCA 2909
Db AAAAGTGATATAAATTTGACCTGTGTAGAACTCTCTGGTATACACTGGCTATTTAAACCA 2580
2910 ATGAGAGGTTTAAATCAACAGAAAACACAGAAATGATCATCAATTTTGTATACCTGCCAT 2969
Db ATGAGAGGTTTAAATCAACAGAAAACACAGAAATGATCATCAATTTTGTATACCTGCCAT 2640
2970 GTAACTACTCTCTGAAATAAATGTGTGCTGATGAGGCTCTACGGTTTGTGGTAGT 3029
Db GTAACTACTCTCTGAAATAAATGTGTGCTGATGAGGCTCTACGGTTTGTGGTAGT 2700
3030 AATCTAATACCTTAAACCCCACTGCTCAAAATCAAAATCAATATCTCTGAGAGACCA 3089
Db AATCTAATACCTTAAACCCCACTGCTCAAAATCAAAATCAATATCTCTGAGAGACCA 2760
3090 GCAATACCATAGAATTTACTAAAAAAA 3120
Db GCAATACCATAGAATTTACTAAAAAAA 2791

RESULT 5

US-09-976-674-8
; Sequence 8, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 4523
; TYPE: DNA

; ORGANISM: Homo sapiens
US-09-976-674-8

Query Match 68.2%; Score 2128; DB 3; Length 4523;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2178; Conservative 0; Mismatches 1; Indels 0

Qy	1	AAGTGTAAAGCCTCGAGGCCAAGCCGCGTCTACTGCGCGCGCTGCTTCTTAGTGCGG	60
Db	1	AAGTGTAAAGCCTCGAGGCCAAGCCGCGTCTACTGCGCGCGCTGCTTCTTAGTGCGG	60
Qy	61	CGTTCGCGCCTGGGTTGTACCGCGCGCGCGCGAGGAAGCACTGCAACACGAGACCG	120
Db	61	CGTTCGCGCCTGGGTTGTACCGCGCGCGCGCGAGGAAGCACTGCAACACGAGACCG	120
Qy	121	GAGTGGAGCGCGCAGCATGAAGCGCGCAGGCCCGCTCCATAGCGCACGTGCGGACGG	180
Db	121	GAGTGGAGCGCGCAGCATGAAGCGCGCAGGCCCGCTCCATAGCGCACGTGCGGACGG	180
Qy	181	TCCGGCGGGCGGGGGAAGAAATGCAACAATGCGCAGCAGCAATGGAACAGAAACAG	240
Db	181	TCCGGCGGGCGGGGGAAGAAATGCAACAATGCGCAGCAGCAATGGAACAGAAACAG	240
Qy	241	CTGGGTGTTGAGATATTTGAAACTCGGACGTGTGAGGAGAAATATGCAATCAGAGATCGG	300
Db	241	CTGGGTGTTGAGATATTTGAAACTCGGACGTGTGAGGAGAAATATGCAATCAGAGATCGG	300
Qy	301	CCTAAATTGGAGCCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAGCTGCTT	360
Db	301	CCTAAATTGGAGCCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAGCTGCTT	360
Qy	361	GCCGATACCCAGAAAATATCATGGCTACATGATGCTAAGGCACCATGATTTTCATGTTT	420
Db	361	GCCGATACCCAGAAAATATCATGGCTACATGATGCTAAGGCACCATGATTTTCATGTTT	420
Qy	421	GTGAAGAGGAATGATCCAGATGGACCTCATTCAGACAGATCTATTACCTTGCCATGCT	480
Db	421	GTGAAGAGGAATGATCCAGATGGACCTCATTCAGACAGATCTATTACCTTGCCATGCT	480
Qy	481	GGTGAGAACAGAGAAAATACACTGTTTTATTCTTGAATTTCCCAAACTATCAATAGACA	540
Db	481	GGTGAGAACAGAGAAAATACACTGTTTTATTCTTGAATTTCCCAAACTATCAATAGACA	540
Qy	541	GCAGTCTTAATGCTCTCTTGGAGCCCTTTTGGATCTTTTTCAGGCAACACTGACAT	600
Db	541	GCAGTCTTAATGCTCTCTTGGAGCCCTTTTGGATCTTTTTCAGGCAACACTGACAT	600
Qy	601	GCAATGTATCTCCGAGAGAGACTATTAAGGAAGAAACCGATTTGGACAGTGGGA	660
Db	601	GCAATGTATCTCCGAGAGAGACTATTAAGGAAGAAACCGATTTGGACAGTGGGA	660
Qy	661	ATTGCGTCTTACGATATACCAAGGAGTGGAACTTTCTGTTTTCAAGCCGGTAGTGGGA	720
Db	661	ATTGCGTCTTACGATATACCAAGGAGTGGAACTTTCTGTTTTCAAGCCGGTAGTGGGA	720
Qy	721	ATTTATCAGCTAAAGATGAGGGCCCAAGGATTTTACGCAACAACCTTTTAAGGCCCAAT	780
Db	721	ATTTATCAGCTAAAGATGAGGGCCCAAGGATTTTACGCAACAACCTTTTAAGGCCCAAT	780
Qy	781	CTAGTGGAACTAGTGTGTCGAACATACGAGTGGATTCAAAAATTATGCCCGCTGATCCA	840
Db	781	CTAGTGGAACTAGTGTGTCGAACATACGAGTGGATTCAAAAATTATGCCCGCTGATCCA	840
Qy	841	GACTGGATGCTTTTATACATAGCAACGATATTTGGATATCTAACATCGTAAACAGAGAA	900
Db	841	GACTGGATGCTTTTATACATAGCAACGATATTTGGATATCTAACATCGTAAACAGAGAA	900
Qy	901	GAAGAGGAGATCTATTATGCAACAATGAGCTAGCCAAACATGGAGAGAGATGCCAGATCA	960
Db	901	GAAGAGGAGATCTATTATGCAACAATGAGCTAGCCAAACATGGAGAGAGATGCCAGATCA	960
Qy	961	GCTGGATGCGCTACCTTGTCTCCAGAGAAATTTGATAGATATTTCTGGCTATTGTGTGG	1020

Db	1621	TTAAGGAAAGCAATATAAAACGATCCAGTGGTGGGCTGCCCTGCTCCAAGTGAATTTCAAG	1680
Qy	1681	TGTCCTATCAAAAGAGGAGATAGCAAAATACCAAGTGGTGAATGGAAAGTTCTTTGGCCGGCAT	1740
Db	1681	TGTCCTATCAAAAGAGGAGATAGCAAAATACCAAGTGGTGAATGGAAAGTTCTTTGGCCGGCAT	1740
Qy	1741	GGATCTAATATCCAAAGTTGATGAAGTCAGAAGCTCGGTATATTTTTGAAGGACCAAAAGAC	1800
Db	1741	GGATCTAATATCCAAAGTTGATGAAGTCAGAAGCTCGGTATATTTTTGAAGGACCAAAAGAC	1800
Qy	1801	TCCCTTTAGAGCATCACCTGTACGTAGTCAGTTAGCTAAATCCTGGAGAGGTGACCAAGG	1860
Db	1801	TCCCTTTAGAGCATCACCTGTACGTAGTCAGTTAGCTAAATCCTGGAGAGGTGACCAAGG	1860
Qy	1861	CTGACTGACCGTGGCTACTCACATTTCTTGCTGCATCAGTCAGCACTGTGACTTCTTTATA	1920
Db	1861	CTGACTGACCGTGGCTACTCACATTTCTTGCTGCATCAGTCAGCACTGTGACTTCTTTATA	1920
Qy	1921	AGTAAGTATAGTAACCAAGAAATCCACACTGTGTCTCCCTTTACAAGCTATCAAGTCCT	1980
Db	1921	AGTAAGTATAGTAACCAAGAAATCCACACTGTGTCTCCCTTTACAAGCTATCAAGTCCT	1980
Qy	1981	GAAGATGACCCAACTTGCAAAACCAAGGAATTTGGGCCAACCATTTTCGGATTCAGCAGGT	2040
Db	1981	GAAGATGACCCAACTTGCAAAACCAAGGAATTTGGGCCAACCATTTTCGGATTCAGCAGGT	2040
Qy	2041	CCTCTTCCTGACTATACCTCTCCAGAAATTTCTCTTTTGAAGTACTACTGGATTTTACA	2100
Db	2041	CCTCTTCCTGACTATACCTCTCCAGAAATTTCTCTTTTGAAGTACTACTGGATTTTACA	2100
Qy	2101	TTGTATGGAGTGCTCTACAAGCCTCATGATCTACAGCCTGGAAAGAAATATCCTACTGTG	2160
Db	2101	TTGTATGGAGTGCTCTACAAGCCTCATGATCTACAGCCTGGAAAGAAATATCCTACTGTG	2160
Qy	2161	CTGTTTCATATATGGTGGTC 2179	
Db	2161	CTGTTTCATATATGGTGGTC 2179	
RESULT 7			
US-09-976-674-22			
; Sequence 22, Application US/09976674			
; Patent No. 6844180			
; GENERAL INFORMATION:			
; APPLICANT: Qi, Steve			
; APPLICANT: Akinsanya, Karen			
; APPLICANT: Riviere, Pierre			
; APPLICANT: Junien, Jean-Louis			
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV			
; FILE REFERENCE: 70669			
; CURRENT APPLICATION NUMBER: US/09/976,674			
; CURRENT FILING DATE: 2001-10-12			
; PRIOR APPLICATION NUMBER: US 60/240,117			
; PRIOR FILING DATE: 2000-10-12			
; NUMBER OF SEQ ID NOS: 61			
; SOFTWARE: PatentIn version 3.1			
; SEQ ID NO 22			
; LENGTH: 4685			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
US-09-976-674-22			
Query Match 63.6%; Score 1984; DB 3; Length 4685;			
Best Local Similarity 99.9%; Pred. No. 0;			
Matches 2034; Conservative 0; Mismatches 1; Indels 0; Gaps 0;			
Qy	1	AAGTCTAAAGCCTCCGAGGCCAAGGCCGTGCTACTGCGCCGCTGCTTCTTAGTCCCG	60
Db	1	AAGTCTAAAGCCTCCGAGGCCAAGGCCGTGCTACTGCGCCGCTGCTTCTTAGTCCCG	60
Qy	61	CGTTCGCCCGCTCGGTGTGTACCGGCCGCCGCCGAGGAAGCCACTGCAACACGAGACCG	120

QY 1201 TCAGAAATATGTTGATGCTGAAGGAGGATCATAGATGTCATAGATAAGCAACTAATT 1260
DB 1201 TCAGAAATATGTTGATGCTGAAGGAGGATCATAGATGTCATAGATAAGCAACTAATT 1260
QY 1261 CAACCTTTTGAGATCTTATTTGAAGGAGTTGAATATATTTGCCAGAGCTGGATGACTCCT 1320
DB 1261 CAACCTTTTGAGATCTTATTTGAAGGAGTTGAATATATTTGCCAGAGCTGGATGACTCCT 1320
QY 1321 GAGGGAATAATGCTTGGTCCATCTCTACTAGATTCGCTCCAGACTCGCCTACAGATAGTG 1380
DB 1321 GAGGGAATAATGCTTGGTCCATCTCTACTAGATTCGCTCCAGACTCGCCTACAGATAGTG 1380
QY 1381 TTGATCTCACTGAATATTTATCCAGTAGAAGATGATGTTATGGAAGGACAGACTC 1440
DB 1381 TTGATCTCACTGAATATTTATCCAGTAGAAGATGATGTTATGGAAGGACAGACTC 1440
QY 1441 ATTGAGTCAGTGCCTGATCTGTGACGCCACTAATTTATCTATGAAGAAACACAGACATC 1500
DB 1441 ATTGAGTCAGTGCCTGATCTGTGACGCCACTAATTTATCTATGAAGAAACACAGACATC 1500
QY 1501 TGGATAAATATCCATGACATCTTTCAATGTTTTTCCCAAAGTCACGAAGAGGAATTTGAG 1560
DB 1501 TGGATAAATATCCATGACATCTTTCAATGTTTTTCCCAAAGTCACGAAGAGGAATTTGAG 1560
QY 1561 TTTATTTTTGCTCTCGAATGCAAAACACAGGTTTTCCGTCATTTATACAAATTTACATCTATT 1620
DB 1561 TTTATTTTTGCTCTCGAATGCAAAACACAGGTTTTCCGTCATTTATACAAATTTACATCTATT 1620
QY 1621 TTAAGGAAAGCAAAATATAACGATCCAGTGTGGGCTGCCTGCTCCAAAGTGTTCGAG 1680
DB 1621 TTAAGGAAAGCAAAATATAACGATCCAGTGTGGGCTGCCTGCTCCAAAGTGTTCGAG 1680
QY 1681 TGCTCTATCAAGAGGAGATAGCAATTTACAGTGTGTGAATGGGAAGTTCTTGGCCGCAT 1740
DB 1681 TGCTCTATCAAGAGGAGATAGCAATTTACAGTGTGTGAATGGGAAGTTCTTGGCCGCAT 1740
QY 1741 GATCTAATATCAAGTTGATGAAGTCAGTAAGCTGAGTAAATCTCTGAGGACCAAAAGAC 1800
DB 1741 GATCTAATATCAAGTTGATGAAGTCAGTAAGCTGAGTAAATCTCTGAGGACCAAAAGAC 1800
QY 1801 TCCCTTTAGAGATCACCTGTACGTAGTCAGTTACGTAAATCTCTGAGAGGTCACAAAG 1860
DB 1801 TCCCTTTAGAGATCACCTGTACGTAGTCAGTTACGTAAATCTCTGAGAGGTCACAAAG 1860
QY 1861 CTGACTGACCGTGGCTACTCACATTTCTGCTGCATCAGTCAGCACTGTGACTTCTTTATA 1920
DB 1861 CTGACTGACCGTGGCTACTCACATTTCTGCTGCATCAGTCAGCACTGTGACTTCTTTATA 1920
QY 1921 AGTAAGTATAGTAACAGAGAAATCCACATCTGTGTCCCTTTACAGCTATCAAGTCCT 1980
DB 1921 AGTAAGTATAGTAACAGAGAAATCCACATCTGTGTCCCTTTACAGCTATCAAGTCCT 1980
QY 1981 GAAGATGACCCAACTTGCRAAAACAAAGGAATTTGGGCCACCATTTTGGATTTCAG 2035
DB 1981 GAAGATGACCCAACTTGCRAAAACAAAGGAATTTGGGCCACCATTTTGGATTTCAG 2035

RESULT 8

US-09-976-674-14
; Sequence 14, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976, 674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12

; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
; LENGTH: 4309
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-976-674-14

Query Match 38.2%; Score 1191; DB 3; Length 4309;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1371; Conservative 0; Mismatches 0; Indels 2; Gaps 1;

QY 1750 ATCCAAGTTGATGAAGTCAGAAAGCTGGTATATTTTGAAGGCACCAAGACTCCCTTTTA 1809
DB 1230 ATCCAAGTTGATGAAGTCAGAAAGCTGGTATATTTTGAAGGCACCAAGACTCCCTTTTA 1289
QY 1810 GAGCATCACCTGTACGTAGTCAGTTAAGTAAATCTCTGGAGAGGTCGCAAGGCTGACTGAC 1869
DB 1290 GAGCATCACCTGTACGTAGTCAGTTAAGTAAATCTCTGGAGAGGTCGCAAGGCTGACTGAC 1349
QY 1870 CGTGGCTACTCACATTTCTTGTCTGATCAGTCAGCACTGTGACTTCTTTATAGTAAGTAT 1929
DB 1350 CGTGGCTACTCACATTTCTTGTCTGATCAGTCAGCACTGTGACTTCTTTATAGTAAGTAT 1409
QY 1930 AGTAACAGAGAAATCCACATCTGTGTCTCTTTTCAAGCTATCAAGTCTCTGAAGATGAC 1989
DB 1410 AGTAACAGAGAAATCCACATCTGTGTCTCTTTTCAAGCTATCAAGTCTCTGAAGATGAC 1469
QY 1990 CCAACTTTGCAAAACAAAGGAATTTTGGGCCACCATTTTGGATTTCAGCAAGGCTCTCTTCC 2049
DB 1470 CCAACTTTGCAAAACAAAGGAATTTTGGGCCACCATTTTGGATTTCAGCAAGGCTCTCTTCC 1529
QY 2050 GACTATACTCTCCAGAAATTTTCTCTTTTGAAGTACTACTTGGATTTCACATTTGATGCG 2109
DB 1530 GACTATACTCTCCAGAAATTTTCTCTTTTGAAGTACTACTTGGATTTCACATTTGATGCG 1589
QY 2110 ATGCTCTACAAGCTCATGATCTACAGCTTGAAGAAATATCTCTACTGTGCTGTTCATA 2169
DB 1590 ATGCTCTACAAGCTCATGATCTACAGCTTGAAGAAATATCTCTACTGTGCTGTTCATA 1649
QY 2170 TATGGTGG--TCCTCAGGTGTCAGTTGGTGAATAAATCGGTTTAAAGGAGTCAAGTATTTCC 2227
DB 1650 TATGGTGGTCTCTCAGGTGTCAGTTGGTGAATAAATCGGTTTAAAGGAGTCAAGTATTTCC 1709
QY 2228 GCTTGAATACCTTAGCCTCTCTAGGTTATGCTGTTGTAGTATGACAAACAGGGGATCCT 2287
DB 1710 GCTTGAATACCTTAGCCTCTCTAGGTTATGCTGTTGTAGTATGACAAACAGGGGATCCT 1769
QY 2288 GTCAACGAGGCTTAAATTTGAAGGCGCTTTTAAATATATAAATGGTCAAAATAGAAATG 2347
DB 1770 GTCAACGAGGCTTAAATTTGAAGGCGCTTTTAAATATATAAATGGTCAAAATAGAAATG 1829
QY 2348 ACAGTACAGTGGAGGAGTCCCAATATCTAGCTCTCGATATGATTTTACATTTAGATC 2407
DB 1830 ACAGTACAGTGGAGGAGTCCCAATATCTAGCTCTCGATATGATTTTACATTTAGATC 1889
QY 2408 GTGTGGGCATCCACGGCTGCTTATGGAGGATACCTCTCCCTGATGCGATTAATGTCAGA 2467
DB 1890 GTGTGGGCATCCACGGCTGCTTATGGAGGATACCTCTCCCTGATGCGATTAATGTCAGA 1949
QY 2468 GGTGAGATATCTTCAGGGTTGCTATTTGCTGGGGCCCCAGTCACTCTGTGGATCTTTCTATG 2527
DB 1950 GGTGAGATATCTTCAGGGTTGCTATTTGCTGGGGCCCCAGTCACTCTGTGGATCTTTCTATG 2009
QY 2528 ATACAGGATACACGGAAACGTTATATGGGTACCCCTGACAGAAATGAAACAGGGCTATTACT 2587
DB 2010 ATACAGGATACACGGAAACGTTATATGGGTACCCCTGACAGAAATGAAACAGGGCTATTACT 2069
QY 2588 TAGGATCTGTGGCCATGCAAGAGAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCT 2647
DB 2070 TAGGATCTGTGGCCATGCAAGAGAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCT 2129
QY 2648 TACATGGTTTCTCGATGAGAAATGTCATTTTGCACATACCAAGTATATTTACTGAGTTTTT 2707


```
Db 2130 TACATGGTTCCTGGATGAGATGTCATTTTGCACATACCGATATATTACTGAGTTTT 2189
Qy 2708 TAGTGAGGGCTGGAAAGCCATATGATTTACAGATCTATCTCCAGAGAGACACAGATAA 2767
Db 2190 TAGTGAGGGCTGGAAAGCCATATGATTTACAGATCTATCTCCAGAGAGACACAGATAA 2249
Qy 2768 GAGTTCCTGAAATCGGAGAAACATATGAACTCGATCTTTTGGACATCTACCTTCAAGAAACC 2827
Db 2250 GAGTTCCTGAAATCGGAGAAACATATGAACTCGATCTTTTGGACATCTACCTTCAAGAAACC 2309
Qy 2828 TTGATCACGATATGCTGCTCTAAAGAGTATATAATTTTGACCTGTGTGAGAACTCTCTGG 2887
Db 2310 TTGATCACGATATGCTGCTCTAAAGAGTATATAATTTTGACCTGTGTGAGAACTCTCTGG 2369
Qy 2888 TATACACTGCTATTTAAACCAATGAGGAGGTTTAAATCAACAGAAACACACAGAAATTGATC 2947
Db 2370 TATACACTGCTATTTAAACCAATGAGGAGGTTTAAATCAACAGAAACACACAGAAATTGATC 2429
Qy 2948 ATACATATTTGATACCTGCCATGTAACATCTACTCTGAAATATAATGTTGGTGCCATGCA 3007
Db 2430 ATACATATTTGATACCTGCCATGTAACATCTACTCTGAAATATAATGTTGGTGCCATGCA 2489
Qy 3008 GGSGTCTACGGTTTGTGGTAGTAATCTAATACCTTTAAACCCACATGCTCAAAATCAAATG 3067
Db 2490 GGSGTCTACGGTTTGTGGTAGTAATCTAATACCTTTAAACCCACATGCTCAAAATCAAATG 2549
Qy 3068 ATACATATTTCTGAGAGACCCAGCAATACCATTAAGAATTTACTTAAAAAAAAAAA 3120
Db 2550 ATACATATTTCTGAGAGACCCAGCAATACCATTAAGAATTTACTTAAAAAAAAAAAA 2602

RESULT 9
US-10-070-464-6
; Sequence 6, Application US/10070464
; Patent No. 684180
; GENERAL INFORMATION:
; APPLICANT: ABBOTT;
; APPLICANT: ABBOTT; Catherine Anne
; FILE OF INVENTION: DIPETIDIL PEPTIDASES
; CURRENT APPLICATION NUMBER: US/10/070,464
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: PCT/AU00/01085
; PRIOR FILING DATE: 2000-09-11
; PRIOR APPLICATION NUMBER: AU PQ5709
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: AU PQ2762
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1669
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-10-070-464-6

Query Match 32.8%; Score 1023; DB 3; Length 1669;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1023; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1164 AACAGGTACAGCAAACTCTAAAGTCACTTTTAAAGATGTCAGAAATATGATGATGCTGA 1223
Db 1 1 AACAGGTACAGCAAACTCTAAAGTCACTTTTAAAGATGTCAGAAATATGATGATGCTGA 60
Qy 1224 AGAAGGATCATAGATGTCATAGATAAGGAATTAATCAACCTTTTGAAGATTTCTATTGA 1283
Db 61 AGAAGGATCATAGATGTCATAGATAAGGAATTAATCAACCTTTTGAAGATTTCTATTGA 120
Qy 1284 AGGAGTGAATATTTGCGAGAGCTGGAGTGGTCTCTGAGGAAATATCTGGTCCAT 1343
Db 121 AGGAGTGAATATTTGCGAGAGCTGGAGTGGTCTCTGAGGAAATATCTGGTCCAT 180
```

```
Qy 1344 CTTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAATTTAT 1403
Db 181 CTTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAATTTAT 240
Qy 1404 CCCAGTGAAGATGATGTTATGGAAGGAGAGACTCATTTAGTGCAGTGCCTGATCTGT 1463
Db 241 CCCAGTGAAGATGATGTTATGGAAGGAGAGACTCATTTAGTGCAGTGCCTGATCTGT 300
Qy 1464 GACGCCCTAATTTATCTATGAAGAAACAAACAGACATCTGGATPAATAATCCATGACATCTT 1523
Db 301 GACGCCCTAATTTATCTATGAAGAAACAAACAGACATCTGGATPAATAATCCATGACATCTT 360
Qy 1524 TCATGTTTTTCCCAAGTACGAGAGGAAATTCAGTTTATTTTGGCTCTGGAATGCAA 1583
Db 361 TCATGTTTTTCCCAAGTACGAGAGGAAATTCAGTTTATTTTGGCTCTGGAATGCAA 420
Qy 1584 AACAGGTTTCCGTCATTTATACAAAATTACATCTATTTTAAAGGAAACAAATATAAACG 1643
Db 421 AACAGGTTTCCGTCATTTATACAAAATTACATCTATTTTAAAGGAAACAAATATAAACG 480
Qy 1644 ATCCAGTGGTGGCTGCTCCTCCAGTATTTCAAGTGTCTCTATCAAAAGAGGAGATAGC 1703
Db 481 ATCCAGTGGTGGCTGCTCCTCCAGTATTTCAAGTGTCTCTATCAAAAGAGGAGATAGC 540
Qy 1704 AATTACCAGTGGTGAATGGGAAGTTCTTGGCCGAGTGGATCTAATATCCCAAGTTGATGA 1763
Db 541 AATTACCAGTGGTGAATGGGAAGTTCTTGGCCGAGTGGATCTAATATCCCAAGTTGATGA 600
Qy 1764 AGTCAGAGGCTGGTATATTTTGAAGGCACAAAGACTCCCTTTAGAGCATCACCTGTA 1823
Db 601 AGTCAGAGGCTGGTATATTTTGAAGGCACAAAGACTCCCTTTAGAGCATCACCTGTA 660
Qy 1824 CGTAGTCAGTTCAGTAAATCCCTGGAGAGTGAACAGGCTGACTGACCTGCTACTCACA 1883
Db 661 CGTAGTCAGTTCAGTAAATCCCTGGAGAGTGAACAGGCTGACTGACCTGCTACTCACA 720
Qy 1884 TTCTTGTCTGATCAGTCAAGCATCTGACCTCTTTATAGTAGTATAGTAAACCAAGAA 1943
Db 721 TTCTTGTCTGATCAGTCAAGCATCTGACCTCTTTATAGTAGTATAGTAAACCAAGAA 780
Qy 1944 TCCACACTGTGTCTCTTTTCAAGCTATCAAGTCTCAAGATGATGCCAATCTGCAAAAC 2003
Db 781 TCCACACTGTGTCTCTTTTCAAGCTATCAAGTCTCAAGATGATGCCAATCTGCAAAAC 840
Qy 2004 AAAGGAATTTTGGGCCACCATTTTGGATTCAGCAGGTCCTCTCTGACTATATCTCTCC 2063
Db 841 AAAGGAATTTTGGGCCACCATTTTGGATTCAGCAGGTCCTCTCTGACTATATCTCTCC 900
Qy 2064 AGAAATTTCTCTTTTGAAGTACTACTGGAATTTACATTTGATGGGATGCTCTCAAGCC 2123
Db 901 AGAAATTTCTCTTTTGAAGTACTACTGGAATTTACATTTGATGGGATGCTCTCAAGCC 960
Qy 2124 TCATGATCTACAGCCTGGAAAGAAATATCTACTGTGTGTTTCATATATGTTGTTCTCA 2183
Db 961 TCATGATCTACAGCCTGGAAAGAAATATCTACTGTGTGTTTCATATATGTTGTTCTCA 1020
Qy 2184 GGT 2186
Db 1021 GGT 1023

RESULT 10
US-09-976-674-10
; Sequence 10, Application US/09976674
; Patent No. 684180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akineanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 20669
; CURRENT APPLICATION NUMBER: US/09/976,674
```


; CURRENT FILING DATE: 2001-10-12									
; PRIOR APPLICATION NUMBER: US 60/240,117									
; PRIOR FILING DATE: 2000-10-12									
; NUMBER OF SEQ ID NOS: 61									
; SOFTWARE: PatentIn version 3.1									
; SEQ ID NO 10									
; LENGTH: 1356									
; TYPE: DNA									
; ORGANISM: Homo sapiens									
US-09-976-674-10									
Query Match 28.1%; Score 877; DB 3; Length 1356;									
Best Local Similarity 99.9%; Pred. No. 0;									
Matches 927; Conservative 0; Mismatches 1; Indels 0; Gaps 0;									
QY	1	AAGTGTAAAGCCTCCGAGGCAAGCGCGTGTCTACTGCGCGCGTCTTCTTAGTGCGG	60						
DB	1	AAGTGTAAAGCCTCCGAGGCAAGCGCGTGTCTACTGCGCGCGTCTTCTTAGTGCGG	60						
QY	61	CGTTCCGCCCTGGGTGTCAACCGCGCGCGCGAGGAGCACCTGCAACCCAGGACCG	120						
DB	61	CGTTCCGCCCTGGGTGTCAACCGCGCGCGCGAGGAGCACCTGCAACCCAGGACCG	120						
QY	121	GAGTGGAGCGCGCAGCATGAAGCGCGCAGGCGCGCTCCATAGCGCAGCTCGGACGG	180						
DB	121	GAGTGGAGCGCGCAGCATGAAGCGCGCAGGCGCGCTCCATAGCGCAGCTCGGACGG	180						
QY	181	TCCGGCGGGCGCGGGGGAAGAAAATGCAACATGGCAGGCAATGGAAAACAGACAG	240						
DB	181	TCCGGCGGGCGCGGGGGAAGAAAATGCAACATGGCAGGCAATGGAAAACAGACAG	240						
QY	241	CTGGGTGTGAGATTTGAACCTGCGGACTGTGAGGAGAAATTTGAATCAGAGGATCG	300						
DB	241	CTGGGTGTGAGATTTGAACCTGCGGACTGTGAGGAGAAATTTGAATCAGAGGATCG	300						
QY	301	CTAAATTGGAGCCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAAAGCTGCT	360						
DB	301	CTAAATTGGAGCCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAAAGCTGCT	360						
QY	361	GCCGATACAGAAAATATCATGGCTACATGATGGCTAAGGACACATGATTTTCATGTT	420						
DB	361	GCCGATACAGAAAATATCATGGCTACATGATGGCTAAGGACACATGATTTTCATGTT	420						
QY	421	GTGAAGAGGAATGATCCAGATGGACTCATTCAGACAGAACTCTATTACTTGGCATGTCT	480						
DB	421	GTGAAGAGGAATGATCCAGATGGACTCATTCAGACAGAACTCTATTACTTGGCATGTCT	480						
QY	481	GGTGAGAAACAGAAAATACACTGTTTTTATCTGAAATTTCCAAAACCTATCAATAGAGCA	540						
DB	481	GGTGAGAAACAGAAAATACACTGTTTTTATCTGAAATTTCCAAAACCTATCAATAGAGCA	540						
QY	541	GCAGTCTTAATGCTCTCTTGGAGGCTCTTTTGGATCTTTTTCAGGCAACACTGGACTAT	600						
DB	541	GCAGTCTTAATGCTCTCTTGGAGGCTCTTTTGGATCTTTTTCAGGCAACACTGGACTAT	600						
QY	601	GGAATGTTTCTCGAGAGAGAACTATTAAGAGAAAGAAAACGCAATGGAAACAGTCGGA	660						
DB	601	GGAATGTTTCTCGAGAGAGAACTATTAAGAGAAAGAAAACGCAATGGAAACAGTCGGA	660						
QY	661	ATTGCTTTTACGATATCACCAAGGAAGTGGAAACATTTCTGTTTCAAGCGGTAGTGA	720						
DB	661	ATTGCTTTTACGATATCACCAAGGAAGTGGAAACATTTCTGTTTCAAGCGGTAGTGA	720						
QY	721	ATTATATCACTAAAGATGAGGGCCACAAAGATTTAGCAACACCTTTTAAGGCCCAAT	780						
DB	721	ATTATATCACTAAAGATGAGGGCCACAAAGATTTAGCAACACCTTTTAAGGCCCAAT	780						
QY	781	CTAGTGAAGAACTAGTGTCCCAACATACGATGGATCCAAAATTTATGCCCGCTGATCCA	840						
DB	781	CTAGTGAAGAACTAGTGTCCCAACATACGATGGATCCAAAATTTATGCCCGCTGATCCA	840						
QY	841	GACTGGATTTGCTTTTATACATAGCAACGATATTTGGATATCTTAACATCGTAACAGAGAA	900						
DB	841	GACTGGATTTGCTTTTATACATAGCAACGATATTTGGATATCTTAACATCGTAACAGAGAA	900						
; Sequence 4, Application US/10070464									
; Patent No. 6881564									
; GENERAL INFORMATION:									
; APPLICANT: ABBOTT, Catherine Anne									
; APPLICANT: CORRELL, Mark Douglas									
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES									
; FILE REFERENCE: GH-007									
; CURRENT APPLICATION NUMBER: US/10/070,464									
; CURRENT FILING DATE: 2002-03-07									
; PRIOR APPLICATION NUMBER: PCT/AU00/01085									
; PRIOR FILING DATE: 2000-09-11									
; PRIOR APPLICATION NUMBER: AU PQ5709									
; PRIOR FILING DATE: 2000-02-18									
; PRIOR APPLICATION NUMBER: AU PQ2762									
; PRIOR FILING DATE: 1999-09-10									
; NUMBER OF SEQ ID NOS: 8									
; SOFTWARE: FastSeq for Windows Version 4.0									
; SEQ ID NO 4									
; LENGTH: 1197									
; TYPE: DNA									
; ORGANISM: Homo Sapiens									
US-10-070-464-4									
Query Match 25.3%; Score 790; DB 3; Length 1197;									
Best Local Similarity 100.0%; Pred. No. 0;									
Matches 790; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	2331	GGGTCAAAATAGAAAATGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGA	2390						
DB	404	GGGTCAAAATAGAAAATGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGA	463						
QY	2391	TTTCATTGATTTAGATCGTGTGGCATCCACGGCTGGTCTTATGGAGGATACCTCTCCCT	2450						
DB	464	TTTCATTGATTTAGATCGTGTGGCATCCACGGCTGGTCTTATGGAGGATACCTCTCCCT	523						
QY	2451	GATGGCATTAAATGACAGAGTTCAGATATCTTCAAGGTTGCTATTTGCTGGGGCCCGAGTCAC	2510						
DB	524	GATGGCATTAAATGACAGAGTTCAGATATCTTCAAGGTTGCTATTTGCTGGGGCCCGAGTCAC	583						
QY	2511	TCTGTGATCTTCTATGATACAGGATACACGGAACTTATATGGTCACTGACCCAGACAGAA	2570						
DB	584	TCTGTGATCTTCTATGATACAGGATACACGGAACTTATATGGTCACTGACCCAGACAGAA	643						
QY	2571	TGACAGGGCTATTACTTAGGATCTGTGACCATGCAAGCAGAAAAGTTCCCTCTGAAACC	2630						
DB	644	TGACAGGGCTATTACTTAGGATCTGTGACCATGCAAGCAGAAAAGTTCCCTCTGAAACC	703						
QY	2631	AAATCGTTTACTGCTCTTACATGGTTTCTGATGAGAAATGTCATTTTGGACATACCAG	2690						
DB	704	AAATCGTTTACTGCTCTTACATGGTTTCTGATGAGAAATGTCATTTTGGACATACCAG	763						
QY	2691	TATATTACTAGTTTTTTTAGTGGGCTGGAAGCCATATGATTTACAGATCTATCTCTCA	2750						
DB	764	TATATTACTAGTTTTTTTAGTGGGCTGGAAGCCATATGATTTACAGATCTATCTCTCA	823						
QY	2751	GGAGACACAGCATTAAGAGTTCTCCTGAATCGGGAGAACATTTATGAATGCAATCTTTTGA	2810						
DB	824	GGAGACACAGCATTAAGAGTTCTCCTGAATCGGGAGAACATTTATGAATGCAATCTTTTGA	883						
QY	2811	CTACCTTCAAGAAAACCTTGGATCAAGTATGCTGCTCTAAAGTGATATAATTTTGACC	2870						
DB	884	CTACCTTCAAGAAAACCTTGGATCAAGTATGCTGCTCTAAAGTGATATAATTTTGACC	943						
QY	2871	TGTGTAGAACTCTCTGATACACTGGCTATTTAACCAATGAGGAGGTTTAAATCAACAG	2930						

Mon May 8 13:41:52 2006

us-10-825-632-2.oligo.rni

Page 16

Db 335 AGAATGTCATTGTCACATACCAGTATATTACTGAGTTTTTTAGTGAGGGCTGGAAGC 394

Search completed: May 3, 2006, 02:28:13
Job time : 940 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 4, 2006, 09:10:06 ; Search time 2192 Seconds
(without alignments)
11770.283 Million cell updates/sec

Title: US-10-825-632-2
Perfect score: 3120
Sequence: 1 aagtgtaagctcgcagg.....agaattactataaaaaaaa 3120

Scoring table: OLIGO NUC
Gapop 60.0, Gapext 60.0

Searched: 9793542 seqs, 4134699005 residues

Word size 8
Total number of hits satisfying chosen parameters: 19584538

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published Applications_NA_Main:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:
3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3120	100.0	3120	7	US-10-415-122-5
2	3120	100.0	3120	8	US-10-825-632-2
3	3003	96.2	3106	7	US-10-311-035-30
4	2957	94.8	3143	6	US-10-170-789-37
5	2889	92.6	4829	3	US-09-976-674-12
6	2889	92.6	4829	3	US-10-982-512-12
7	2547	81.6	2649	6	US-10-170-789-39
8	2496	80.0	2649	6	US-10-054-776-1
9	2495	80.0	2671	3	US-09-976-674-2
10	2495	80.0	2671	9	US-10-982-512-2
11	2135	68.4	2830	9	US-10-956-157-2177
12	2128	68.2	4523	3	US-09-976-674-8
13	2128	68.2	4523	9	US-10-982-512-8
14	2128	68.2	4676	3	US-09-976-674-20
15	2128	68.2	4676	3	US-10-982-512-20
16	1984	63.6	4685	3	US-09-976-674-22
17	1984	63.6	4685	9	US-10-982-512-22
18	1191	38.2	4309	3	US-09-976-674-14
19	1191	38.2	4309	9	US-10-982-512-14
20	1164	37.3	2510	7	US-10-275-505-16
21	1164	37.3	2510	10	US-11-140-224-16
22	1023	32.8	1669	8	US-10-825-632-6
23	877	28.1	1356	3	US-09-976-674-10

24	877	28.1	1356	9	US-10-982-512-10	Sequence 10, Appl
25	790	25.3	1197	8	US-10-825-632-4	Sequence 4, Appl
26	789	25.3	1083	8	US-10-825-632-8	Sequence 8, Appl
27	760	24.4	832	3	US-09-976-674-18	Sequence 18, Appl
28	760	24.4	832	3	US-10-982-512-18	Sequence 18, Appl
29	737	23.6	925	6	US-10-264-237-710	Sequence 710, App
30	590	18.9	600	9	US-10-956-157-7412	Sequence 7412, Ap
31	472	15.1	620	3	US-09-976-674-16	Sequence 16, Appl
32	472	15.1	620	9	US-10-982-512-16	Sequence 16, Appl
33	431	13.8	502	3	US-09-918-995-19585	Sequence 19585, A
34	254	8.1	561	3	US-09-764-891-877	Sequence 877, App
C 35	167	5.4	168	9	US-10-467-851-254	Sequence 254, App
C 36	156	5.0	4797	3	US-09-764-891-7074	Sequence 7074, Ap
C 37	118	3.8	1346	4	US-09-925-065A-669313	Sequence 669313,
C 38	112	3.6	609	4	US-09-925-065A-743558	Sequence 743558,
C 39	112	3.6	609	4	US-09-925-065A-743559	Sequence 743559,
C 40	73	2.3	612	4	US-09-925-065A-818446	Sequence 818446,
C 41	46	1.5	60	3	US-09-908-975-15085	Sequence 15085, A
C 42	30	1.0	30	8	US-10-825-632-27	Sequence 27, Appl
C 43	28	0.9	497	4	US-09-925-065A-482467	Sequence 482467,
44	27	0.9	392	3	US-09-783-590-879	Sequence 879, App
45	25	0.8	25	9	US-10-956-157-58438	Sequence 58438, A

ALIGNMENTS

RESULT 1
US-10-415-122-5
; Sequence 5, Application US/10415122
; Publication No. US20040053369A1
; GENERAL INFORMATION:
; APPLICANT: THE UNIVERSITY OF SYDNEY
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES
; FILE REFERENCE: FPI5217
; CURRENT APPLICATION NUMBER: US/10/415,122
; CURRENT FILING DATE: 2003-08-07
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 5
; LENGTH: 3120
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-415-122-5

Query Match		100.0%;	Score 3120;	DB 7;	Length 3120;
Best Local Similarity		100.0%;	Pred. No. 0;		
Matches 3120;		Conservative	0;	Mismatches	0;
				Indels	0;
				Gaps	0;
Qy	1	AAGTGCTAAAGCCTCCGAGGCCAAGCCGCTGCTACTGCGCGCGCTGCTTCTTAGTGC	60		
Db	1	AAGTGCTAAAGCCTCCGAGGCCAAGCCGCTGCTACTGCGCGCGCTGCTTCTTAGTGC	60		
Qy	61	CGTTCGCGCGCTGGGTGTGTCACCGCGCGCGCGCGAGGAGCCACTCCACCCAGGACCG	120		
Db	61	CGTTCGCGCGCTGGGTGTGTCACCGCGCGCGCGAGGAGCCACTCCACCCAGGACCG	120		
Qy	121	GAGTGGAGCGCGCGCAGCATGAAGCGCGCGCGCTCCATAGCGCAGCATGCTGGGACCG	180		
Db	121	GAGTGGAGCGCGCGCAGCATGAAGCGCGCGCGCTCCATAGCGCAGCATGCTGGGACCG	180		
Qy	181	TCCGGCGCGCGCGCGGGGGAAGAAAATCAACATGGCAGCAGCAATGGAACAGAACAG	240		
Db	181	TCCGGCGCGCGCGCGGGGGAAGAAAATCAACATGGCAGCAGCAATGGAACAGAACAG	240		
Qy	241	CTGGGTGTGAGATATTGAACTGCGGACTGTGAGGAGAAATTTGAATACAGGATCGG	300		
Db	241	CTGGGTGTGAGATATTGAACTGCGGACTGTGAGGAGAAATTTGAATACAGGATCGG	300		
Qy	301	CCTAAATTGAGCGCTTTTATGTTGAGCGGTATTCCTGAGTTCAGCTTAAAGTCTT	360		
Db	301	CCTAAATTGAGCGCTTTTATGTTGAGCGGTATTCCTGAGTTCAGCTTAAAGTCTT	360		

QY 361 GCCGATACCGAAAAATATCATGGCTACATGATGGCTTAAGGCACACCATGATTTTCATGTTT 420
Db 361 GCCGATACCGAAAAATATCATGGCTACATGATGGCTTAAGGCACACCATGATTTTCATGTTT 420
QY 421 GTGAAGAGGATGATCCAGATGCACTCATTCAGACAGATCTATTACCTTGCCTGCTCT 480
Db 421 GTGAAGAGGATGATCCAGATGCACTCATTCAGACAGATCTATTACCTTGCCTGCTCT 480
QY 481 GGTGAGACACAGAGAAATACACTGTTTTTATCTGAAATTCGAAATTCGAAATTCATCAATAGAGCA 540
Db 481 GGTGAGACACAGAGAAATACACTGTTTTTATCTGAAATTCGAAATTCGAAATTCATCAATAGAGCA 540
QY 541 GCAGCTTAATGCTCTCTTGGAAAGCCTCTTTTGGATCTTTTTCAGGCAACACTGGACTAT 600
Db 541 GCAGCTTAATGCTCTCTTGGAAAGCCTCTTTTGGATCTTTTTCAGGCAACACTGGACTAT 600
QY 601 GGAATGTATTTCTCGAGAAGAGAACTATTAAAGAGAGAGAAACGCAATTTGGAAACAGTCGGA 660
Db 601 GGAATGTATTTCTCGAGAAGAGAACTATTAAAGAGAGAGAAACGCAATTTGGAAACAGTCGGA 660
QY 661 ATTGCTTTCTTACGATTATACCAAGGAAGTGGAAATTTCTGTTTCAAGCCGGTAGTGA 720
Db 661 ATTGCTTTCTTACGATTATACCAAGGAAGTGGAAATTTCTGTTTCAAGCCGGTAGTGA 720
QY 721 ATTTATCACGTAAAGATGAGGCCACAAAGGATTTACGCAACACCTTTTAAAGCCCAAT 780
Db 721 ATTTATCACGTAAAGATGAGGCCACAAAGGATTTACGCAACACCTTTTAAAGCCCAAT 780
QY 781 CTAGTGGAACTAGTTGTCCCAACATACGATGGATCCAAATTTATCCCGCTGATCCA 840
Db 781 CTAGTGGAACTAGTTGTCCCAACATACGATGGATCCAAATTTATCCCGCTGATCCA 840
QY 841 GACTGGATTCCTTTTATACATAGCAACGATTTTGGATATCTTAACATCGTAAACAGAGAA 900
Db 841 GACTGGATTCCTTTTATACATAGCAACGATTTTGGATATCTTAACATCGTAAACAGAGAA 900
QY 901 GAAAGGAGACTCATTTATGTGCAAAATGAGCTAGCCAAATGGAAGAGATGCCAGATCA 960
Db 901 GAAAGGAGACTCATTTATGTGCAAAATGAGCTAGCCAAATGGAAGAGATGCCAGATCA 960
QY 961 GCTGGAGTCTGCTACCTTTGTTCTCAAGAGAAATTTGATAGATATTCGCTATTTGGTGG 1020
Db 961 GCTGGAGTCTGCTACCTTTGTTCTCAAGAGAAATTTGATAGATATTCGCTATTTGGTGG 1020
QY 1021 TGTCCAAAAGCTGAAACAACTCCAGTGGTGGTAAATTTCTTAGAATTTCTATAGAGAA 1080
Db 1021 TGTCCAAAAGCTGAAACAACTCCAGTGGTGGTAAATTTCTTAGAATTTCTATAGAGAA 1080
QY 1081 AATGATGAATCTGAGGTGAAATTTATTCATGTTACATCCCTATGTTGGAAACAAAGGAGG 1140
Db 1081 AATGATGAATCTGAGGTGAAATTTATTCATGTTACATCCCTATGTTGGAAACAAAGGAGG 1140
QY 1141 GCAGATTCATTCCGTTTATCCTTAAACAGGTACAGCAAAATCCTTAAAGTCACCTTTTAAAGATG 1200
Db 1141 GCAGATTCATTCCGTTTATCCTTAAACAGGTACAGCAAAATCCTTAAAGTCACCTTTTAAAGATG 1200
QY 1201 TCAGAAATATGATTGATCTCGAAGGAGGATCATAGATGTCATAGATTAAGGAACTAAT 1260
Db 1201 TCAGAAATATGATTGATCTCGAAGGAGGATCATAGATGTCATAGATTAAGGAACTAAT 1260
QY 1261 CAACCTTTTGAGATTCATTTTGAAGAGTTCGATATATTGCGAGCTGATGCACTCT 1320
Db 1261 CAACCTTTTGAGATTCATTTTGAAGAGTTCGATATATTGCGAGCTGATGCACTCT 1320
QY 1321 GAGGAAATATGCTTGGTCCATCTCATAGATCGCTCCAGACTCCGCTTACAGATAGTG 1380
Db 1321 GAGGAAATATGCTTGGTCCATCTCATAGATCGCTCCAGACTCCGCTTACAGATAGTG 1380
QY 1381 TTGATCTCACCTGAATTTATTTATCCAGTAGAAGATGATGTTATGGAAGGCGAGACTC 1440
Db 1381 TTGATCTCACCTGAATTTATTTATCCAGTAGAAGATGATGTTATGGAAGGCGAGACTC 1440
QY 1441 ATTGAGTCAGTGCCTGATTTCTGTGACGCCCACTAATTTATCTATGAAGAAACACAGACATC 1500

Db 1441 ATTGAGTCAGTGCCTGATTTCTGTGACGCCCACTAATTTATCTATGAAGAAACACAGACATC 1500
QY 1501 TGGATAAATATCCATGACATCTTTTTCATGTTTTCCTCCAAAGTCACGAGAGAGAAATTTGAG 1560
Db 1501 TGGATAAATATCCATGACATCTTTTTCATGTTTTCCTCCAAAGTCACGAGAGAGAAATTTGAG 1560
QY 1561 TTTATTTTTCCTCTGAAATGCAAAACAGGTTTCCGTCATTTTATACAAAATTTACATCTATT 1620
Db 1561 TTTATTTTTCCTCTGAAATGCAAAACAGGTTTCCGTCATTTTATACAAAATTTACATCTATT 1620
QY 1621 TTTAAAGGAAAGCAAAATATAAACGATCCAGTGTGGGCTGCCTGCTCCCAAGTGAATTTCAAG 1680
Db 1621 TTTAAAGGAAAGCAAAATATAAACGATCCAGTGTGGGCTGCCTGCTCCCAAGTGAATTTCAAG 1680
QY 1681 TGTCTTATCAAGAGAGAGATAGCAATTTACAGTGTGTAATGGGAAGTTCTTTGGCCGGCAT 1740
Db 1681 TGTCTTATCAAGAGAGAGATAGCAATTTACAGTGTGTAATGGGAAGTTCTTTGGCCGGCAT 1740
QY 1741 GGATCTAATATCCAAGTTGATGAAGTCAGAAGGCTGGTATATTTTGAAGGCCACCAAGAC 1800
Db 1741 GGATCTAATATCCAAGTTGATGAAGTCAGAAGGCTGGTATATTTTGAAGGCCACCAAGAC 1800
QY 1801 TCCCCTTTTAGACATCACCTGTACGTAGTCAGTTACGTAAATCTCGAGAGGTGACAAGG 1860
Db 1801 TCCCCTTTTAGACATCACCTGTACGTAGTCAGTTACGTAAATCTCGAGAGGTGACAAGG 1860
QY 1861 CTGACTGACCGTGGCTACTCACAATTTCTGCTGCATCAGTCAGCACTGTGACTTCTTTATA 1920
Db 1861 CTGACTGACCGTGGCTACTCACAATTTCTGCTGCATCAGTCAGCACTGTGACTTCTTTATA 1920
QY 1921 AGTAAGTATAGTAAACAGAGAAATCCACATGTGTGTCCTTTTCAAGCTATCAAGTCCT 1980
Db 1921 AGTAAGTATAGTAAACAGAGAAATCCACATGTGTGTCCTTTTCAAGCTATCAAGTCCT 1980
QY 1981 GAAGATGACCCAACTTTCGAAAAACAAAGGAATTTTGGGCCACCATTTTGGATTCAGCAGGT 2040
Db 1981 GAAGATGACCCAACTTTCGAAAAACAAAGGAATTTTGGGCCACCATTTTGGATTCAGCAGGT 2040
QY 2041 CCTCTTCTGACTATCTCCTCCAGAAATTTTCTCTTTTGAAGTACTACTTGGATTTTACA 2100
Db 2041 CCTCTTCTGACTATCTCCTCCAGAAATTTTCTCTTTTGAAGTACTACTTGGATTTTACA 2100
QY 2101 TTGTATGGGATGCTCTACAGGCTCATGATCTACAGCCTGGAAAGAAATATCTTACTGTG 2160
Db 2101 TTGTATGGGATGCTCTACAGGCTCATGATCTACAGCCTGGAAAGAAATATCTTACTGTG 2160
QY 2161 CTGTTTATATATGTTGGTCTCTCAGGTGAGTTGGTGAATAATCGGTTTAAAGGAGTCAAG 2220
Db 2161 CTGTTTATATATGTTGGTCTCTCAGGTGAGTTGGTGAATAATCGGTTTAAAGGAGTCAAG 2220
QY 2221 TATTTCCGCTTGAATACCTAGCCTCTCTAGGTTATGTTGTAGTAGAGACAACAGG 2280
Db 2221 TATTTCCGCTTGAATACCTAGCCTCTCTAGGTTATGTTGTAGTAGAGACAACAGG 2280
QY 2281 GGATCTGTACCGAGGCTTAAATTTGAAGGCGCTTAAATATAAATGCGGTCAATA 2340
Db 2281 GGATCTGTACCGAGGCTTAAATTTGAAGGCGCTTAAATATAAATGCGGTCAATA 2340
QY 2341 GAAATTTGACGATCAGGTGGAAGACTCCAATATCTAGCTTCTCGATATGATTTTCAATGAC 2400
Db 2341 GAAATTTGACGATCAGGTGGAAGACTCCAATATCTAGCTTCTCGATATGATTTTCAATGAC 2400
QY 2401 TTAGATCTGTGGGATCCAGGCTGTCTTATGGAGGATACCTCTCCCTGATGGCAATTA 2460
Db 2401 TTAGATCTGTGGGATCCAGGCTGTCTTATGGAGGATACCTCTCCCTGATGGCAATTA 2460
QY 2461 ATCCAGAGGTTCAGATATCTTCAAGGTTGCTATTTCTGGGCCCCAGTCCTCTGTGGATC 2520
Db 2461 ATCCAGAGGTTCAGATATCTTCAAGGTTGCTATTTCTGGGCCCCAGTCCTCTGTGGATC 2520
QY 2521 TTCTATGATACAGGATACAGGAACTGTTATATGGGTCACTCTGACCAAGATGAACAGGCG 2580

QY 1141 GCAGATTCATTCCGTTATCTCTAAACAGGTACAGCAATCCCTAAAGTCACCTTTTAAGATG 1200
Db 1141 GCAGATTCATTCCGTTATCTCTAAACAGGTACAGCAATCCCTAAAGTCACCTTTTAAGATG 1200
QY 1201 TCAGAAATAATGATGTGATGCTGAAGGAAGATCATAGATGTCATAGATAAGAACTAAAT 1260
Db 1201 TCAGAAATAATGATGTGATGCTGAAGGAAGATCATAGATGTCATAGATAAGAACTAAAT 1260
QY 1261 CAACCTTTTGAGATCTTATTTGAAGGAGTTGAATATATTGCGAGAGCTGGATGACCTCT 1320
Db 1261 CAACCTTTTGAGATCTTATTTGAAGGAGTTGAATATATTGCGAGAGCTGGATGACCTCT 1320
QY 1321 GAGGGAATAATGCTTGGTCCATCTCTACTAGATCGCTCCGACCTCGCTACAGATAGTG 1380
Db 1321 GAGGGAATAATGCTTGGTCCATCTCTACTAGATCGCTCCGACCTCGCTACAGATAGTG 1380
QY 1381 TTGATCTCACTGAAATTAATTTATCCAGTAGAAGATGATGTTATGGAAGGCAGAGACTC 1440
Db 1381 TTGATCTCACTGAAATTAATTTATCCAGTAGAAGATGATGTTATGGAAGGCAGAGACTC 1440
QY 1441 ATTCAGTCAGTGCCTGATCTGTGAGCCCACTAAATATCTATGAAGAACACAGACATC 1500
Db 1441 ATTCAGTCAGTGCCTGATCTGTGAGCCCACTAAATATCTATGAAGAACACAGACATC 1500
QY 1501 TGGATAAATATCCATGACATCTTTCATGTTTCCCAAGTCCAGAGAGGAATTCAG 1560
Db 1501 TGGATAAATATCCATGACATCTTTCATGTTTCCCAAGTCCAGAGAGGAATTCAG 1560
QY 1561 TTTATTTTGGCTCTGAAATCAAAACAGGTTTCGGTCACTTTATACAAATTAATCATCTATT 1620
Db 1561 TTTATTTTGGCTCTGAAATCAAAACAGGTTTCGGTCACTTTATACAAATTAATCATCTATT 1620
QY 1621 TTAAGGAAAGCAAAATATAACGATCCAGTGTGGCTGCCCTCCCAAGTGATTTCAAG 1680
Db 1621 TTAAGGAAAGCAAAATATAACGATCCAGTGTGGCTGCCCTCCCAAGTGATTTCAAG 1680
QY 1681 TGTCTCTATCAAAGGAGATAGCAATTTACAGTGTGTGAATGGGAAGTTCTTTGGCCGGCAT 1740
Db 1681 TGTCTCTATCAAAGGAGATAGCAATTTACAGTGTGTGAATGGGAAGTTCTTTGGCCGGCAT 1740
QY 1741 GGATCTAATATCCAAAGTTGATGAAGTCAGAAAGCTGGTATATTTTGAAGGCACCAAGAGAC 1800
Db 1741 GGATCTAATATCCAAAGTTGATGAAGTCAGAAAGCTGGTATATTTTGAAGGCACCAAGAGAC 1800
QY 1801 TCCCTTTTAGAGCATCACCTGTAGTCAGTTACGTAAATCTCTGGAGAGGTGACAAGG 1860
Db 1801 TCCCTTTTAGAGCATCACCTGTAGTCAGTTACGTAAATCTCTGGAGAGGTGACAAGG 1860
QY 1861 CTGACTGACCGTGGCTACTCACATTTCTGCTGCATCAGTCAGCACTGTGACTCTTTTATA 1920
Db 1861 CTGACTGACCGTGGCTACTCACATTTCTGCTGCATCAGTCAGCACTGTGACTCTTTTATA 1920
QY 1921 AGTAAGTATAGTAACCAAGAAATCCACATGTGTGTCCTTTACAGCTATCAAGTCCT 1980
Db 1921 AGTAAGTATAGTAACCAAGAAATCCACATGTGTGTCCTTTTACAGCTATCAAGTCCT 1980
QY 1981 GAAGATGACCCAACTTGCAAAAACAAAGGAATTTTGGGCCACATTTTGGATTCAGCAGGT 2040
Db 1981 GAAGATGACCCAACTTGCAAAAACAAAGGAATTTTGGGCCACATTTTGGATTCAGCAGGT 2040
QY 2041 CCTCTTCTGACTATCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGAATTTTACA 2100
Db 2041 CCTCTTCTGACTATCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGAATTTTACA 2100
QY 2101 TTGATGGGATGCTCTACAAGCCTCATGATCTACAGCCTGGAAAGAAATATCTTACTGTG 2160
Db 2101 TTGATGGGATGCTCTACAAGCCTCATGATCTACAGCCTGGAAAGAAATATCTTACTGTG 2160
QY 2161 CTGTTTCATATATGGTGGTCTCAGGTGCGAGTTGGTGAATATTCGGTTTAAAGAGGTCAAG 2220
Db 2161 CTGTTTCATATATGGTGGTCTCAGGTGCGAGTTGGTGAATATTCGGTTTAAAGAGGTCAAG 2220
QY 2221 TATTTCCGCTTGAATACCCCTAGCCTCTCTAGGTTATGTGTTAGTGTATAGACACAGG 2280

Db 2221 TATTTCCGCTTGAATACCCCTAGCCTCTCTAGGTTATGTGTTAGTGTATAGACACAGG 2280
QY 2281 GGATCTCTGTACCGAGGGCTTAAATTTTGAAGGGCCCTTTAAATATAAAATGGGTCAATA 2340
Db 2281 GGATCTCTGTACCGAGGGCTTAAATTTTGAAGGGCCCTTTAAATATAAAATGGGTCAATA 2340
QY 2341 GAAATTCAGCATCAGGTGGAGGACCTCCAATATCTAGCTTCTCGATATGATTTTCATTGAC 2400
Db 2341 GAAATTCAGCATCAGGTGGAGGACCTCCAATATCTAGCTTCTCGATATGATTTTCATTGAC 2400
QY 2401 TTAGATCGTGTGGGCATCCACGGCTGGTCTTATGAGGAGATACCTCTCCCTGATGGCATTA 2460
Db 2401 TTAGATCGTGTGGGCATCCACGGCTGGTCTTATGAGGAGATACCTCTCCCTGATGGCATTA 2460
QY 2461 ATGCAGAGGTGAGATATCTTACGGGTTGCTATTTGCTGGGGCCCAAGTCCTCTGTGGATC 2520
Db 2461 ATGCAGAGGTGAGATATCTTACGGGTTGCTATTTGCTGGGGCCCAAGTCCTCTGTGGATC 2520
QY 2521 TTCTATGATACAGGATACAGGAACGTTATATGAGGTCAACCCTGACCAAGATGAACAGGC 2580
Db 2521 TTCTATGATACAGGATACAGGAACGTTATATGAGGTCAACCCTGACCAAGATGAACAGGC 2580
QY 2581 TATTACTTAGATCTGTGGCCATGCAAGCAGAAAGTTCCCTCTGAACCAATCGTTTA 2640
Db 2581 TATTACTTAGATCTGTGGCCATGCAAGCAGAAAGTTCCCTCTGAACCAATCGTTTA 2640
QY 2641 CTGCTCTTACATGGTTTCTCGATGAGAAATGTCATTTTGCATACCAAGTATATTACTG 2700
Db 2641 CTGCTCTTACATGGTTTCTCGATGAGAAATGTCATTTTGCATACCAAGTATATTACTG 2700
QY 2701 AGTTTTTATGAGGGCTGGAAAGCCATATGATTTACAGATCTATCTCTCAGAGAGACAC 2760
Db 2701 AGTTTTTATGAGGGCTGGAAAGCCATATGATTTTACAGATCTATCTCTCAGAGAGACAC 2760
QY 2761 AGCATAAGAGTCTCTGAATCGGAGAACATTAATGAACATCTTTTGCACCTACCTTCAA 2820
Db 2761 AGCATAAGAGTCTCTGAATCGGAGAACATTAATGAACATCTTTTGCACCTACCTTCAA 2820
QY 2821 GAAACCTTGGATCAGTATTGCTGCTCTAAAAGTGATATAATTTTGACCTGTGTAGAAC 2880
Db 2821 GAAACCTTGGATCAGTATTGCTGCTCTAAAAGTGATATAATTTTGACCTGTGTAGAAC 2880
QY 2881 TCTCTGTGTATACACTGGCTATTTAAACAAATGAGGAGGTTTAAATCAACAGAAAAACACAGA 2940
Db 2881 TCTCTGTGTATACACTGGCTATTTAAACAAATGAGGAGGTTTAAATCAACAGAAAAACACAGA 2940
QY 2941 ATTGATCATCACATTTTGTATACCTGCAATGTAACATCTACTCTGAAAAATAATGTGTG 3000
Db 2941 ATTGATCATCACATTTTGTATACCTGCAATGTAACATCTACTCTGAAAAATAATGTGTG 3000
QY 3001 CCATGCAAGGCTCTACGGTTTGTGTAGTAATCTAATACCTTAACCCACATGCTCAAAA 3060
Db 3001 CCATGCAAGGCTCTACGGTTTGTGTAGTAATCTAATACCTTAACCCACATGCTCAAAA 3060
QY 3061 TCAATGATACATATTCTCTGAGAGACCCAGCAATACCATAGAAATTTACTAAAAAANA 3120
Db 3061 TCAATGATACATATTCTCTGAGAGACCCAGCAATACCATAGAAATTTACTAAAAAANA 3120

RESULT 3

US-10-311-035-30
; Sequence 30, Application US/10311035
; Publication No. US20040023243A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: YUE, Henry
; APPLICANT: ELLIOTT, Vicki
; APPLICANT: GANDHI, Ameena R.
; APPLICANT: LAL, Preeti
; APPLICANT: AU-YOUNG, Janice
; APPLICANT: TRIBOULEY, Catherine M.
; APPLICANT: DELEGANE, Angelo M.

APPLICANT: BAUGHN, Mariah R.	Db		482	GAATAATACACTGTTTATTTCTGAATTCGCAAAATCATCATAGAGCAGCTCTTAATG	541
APPLICANT: NGUYEN, Dannel B.	Qy		553	CTCTCTTTGGAAGCCTCTTTTGGATCTTTTCAGGCAACACTGGACTATGGAATGTATCT	612
APPLICANT: LEE, Ernestine A.	Db		542	CTCTCTTTGGAAGCCTCTTTTGGATCTTTTCAGGCAACACTGGACTATGGAATGTATCT	601
APPLICANT: KHAN, Farrah A.	Qy		613	CGAGAAGAACTATATAGAGAAGAAACCATTTGGACAGTCGGATTCCTCTTAC	672
APPLICANT: CHAWLA, Narinder K.	Db		602	CGAGAAGAACTATATAGAGAAGAAACCATTTGGACAGTCGGATTCCTCTTAC	661
APPLICANT: YAO, Monique G.	Qy		673	GATTATCAACCAAGGAGTGAACATTTCTGTTTCAAGCGGTAGTGGAAATTTATCAAGTA	732
APPLICANT: LU, Dyung Aina M.	Db		662	GATTATCAACCAAGGAGTGAACATTTCTGTTTCAAGCGGTAGTGGAAATTTATCAAGTA	721
APPLICANT: ARVIZU, Chandra S.	Qy		733	AAAGATGGAGGGCCACAAGGATTTACGCAACAACTTTAAGGCCCAATCTCTAGTGGAACT	792
APPLICANT: TANG, Y. Tom	Db		722	AAAGATGGAGGGCCACAAGGATTTACGCAACAACTTTAAGGCCCAATCTCTAGTGGAACT	781
APPLICANT: WALSH, Roderick T.	Qy		793	AGTTGTCCCAACATACGGATGGATCCAAATTTATGCCCGCTGATCCAGACTGGATTGCT	852
APPLICANT: LU, Yan	Db		782	AGTTGTCCCAACATACGGATGGATCCAAATTTATGCCCGCTGATCCAGACTGGATTGCT	841
APPLICANT: RAMKUMAR, Jayalaximi	Qy		853	TTTATACATAGCAACGATATTTGGATATCTAACTATCTAACTGTAACCCAGAGAAGAAAGAGACTC	912
APPLICANT: REDDY, Roopa	Db		842	TTTATACATAGCAACGATATTTGGATATCTAACTATCTAACTGTAACCCAGAGAAGAAAGAGACTC	901
APPLICANT: DAS, Depodriya	Qy		913	ACTTATGTGCACAAATGAGCTAGCCAAACATGGAAGAAGATGCCAGATCAGCTGGAGTCGCT	972
APPLICANT: KEARNEY, Liam	Db		902	ACTTATGTGCACAAATGAGCTAGCCAAACATGGAAGAAGATGCCAGATCAGCTGGAGTCGCT	961
APPLICANT: KALLICK, Deborah A.	Qy		973	ACCTTTGTTCTCCAAAGAAGATTTGATAGATATCTGCGTATTTGGTGTGTCCTCAAAAGCT	1032
TITLE OF INVENTION: Proteases	Db		962	ACCTTTGTTCTCCAAAGAAGATTTGATAGATATTTGGTGTGTCCTCAAAAGCT	1021
FILE REFERENCE: PI-0123 PCT	Qy		1033	GAAACAACTCCAGTGGTGTAAATTTCTAGAAATTTCTATATGAAGAAATGATGAATCT	1092
CURRENT APPLICATION NUMBER: US/10/311,035	Db		1022	GAAACAACTCCAGTGGTGTAAATTTCTAGAAATTTCTATATGAAGAAATGATGAATCT	1081
CURRENT FILING DATE: 2002-12-10	Qy		1093	GAGGTGGAATTTATTCATGTTACATCCCTATTTGGAAACAAGAGGCGAGATTCATTC	1152
PRIOR APPLICATION NUMBER: 60/212,336; 60/213,955; 60/215,396; 60/216,821; 60/218,946	Db		1082	GAGGTGGAATTTATTCATGTTACATCCCTATTTGGAAACAAGAGGCGAGATTCATTC	1141
PRIOR FILING DATE: 2000-06-16; 2000-06-22; 2000-06-29; 2000-07-07; 2000-07-14	Qy		1153	CGTTATCTTAAACACAGGTACAGCAATCTTAAAGTCACTTTTAAGATGTCAAGATAATG	1212
NUMBER OF SEQ ID NOS: 42	Db		1142	CGTTATCTTAAACACAGGTACAGCAATCTTAAAGTCACTTTTAAGATGTCAAGATAATG	1201
SOFTWARE: PERL Program	Qy		1213	ATTGATCTCAAGGAAGGATCATAGATGTCATAGATAGAGAACTAAATTCACCTTTTGAG	1272
SEQ ID NO 30	Db		1202	ATTGATCTCAAGGAAGGATCATAGATGTCATAGATAGAGAACTAAATTCACCTTTTGAG	1261
LENGTH: 3106	Qy		1273	ATTCTATTTCGAAGAGTTCGAATATATTGGCAGAGCTGGATGGACTCTCTGAGGGAATAAT	1332
TYPE: DNA	Db		1262	ATTCTATTTCGAAGAGTTCGAATATATTGGCAGAGCTGGATGGACTCTCTGAGGGAATAAT	1321
ORGANISM: Homo sapiens	Qy		1333	CGTTGGTCCATCTTACATAGATCGCTCCAGACTCGCGCTACAGATAGTGTGATCTCACCT	1392
FEATURE:	Db		1322	CGTTGGTCCATCTTACATAGATCGCTCCAGACTCGCGCTACAGATAGTGTGATCTCACCT	1381
NAME/KEY: misc feature	Qy		1393	GAAATATTTATCCAGTAGAAGATGATGTTATGGAAGGCAGAGACTCATTTGAGTCAGTG	1452
OTHER INFORMATION: Incyte ID No. US20040023243A1 7160544CB1	Db		1382	GAAATATTTATCCAGTAGAAGATGATGTTATGGAAGGCAGAGACTCATTTGAGTCAGTG	1441
US-10-311-035-30	Qy		1453	CCTGATTTCTGTGAGCGCCACTTAAATTTATCTATGAAGAAACAACAGACATCTGGATAAATATC	1512
Query Match : 96.2%; Score 3003; DB 7; Length 3106;	Db		1442	CCTGATTTCTGTGAGCGCCACTTAAATTTATCTATGAAGAAACAACAGACATCTGGATAAATATC	1501
Best Local Similarity 99.9%; Pred. No. 0;	Qy		1513	CATGACATCTTTCATGTTTTTCCCAAAGTCAAGAGAGAAATTTGAGTTATTTTTCGCC	1572
Matches 3103; Conservative 0; Mismatches 2; Indels 0; Gaps 0;	Db		1502	CATGACATCTTTCATGTTTTTCCCAAAGTCAAGAGAGAAATTTGAGTTATTTTTCGCC	1561
13 CTCGAGGCCAAGCGCGCTGCTACTGCGCGCGCTGCTTCTTAGTGCAGCTTCCCGCCT 72	Qy		1573	TCTGAATGCAAAACAGGTTTCCGTCATTTATATCAAAATTTACATCTATTTTAAAGGAAAGC	1632
2 CTCGAGGCCAAGCGCGCTGCTACTGCGCGCGCTGCTTCTTAGTGCAGCTTCCCGCCT 61	Db				
73 GGGTTGTACCGCGCGCGCGCGAGGAAGCCACTGCAACACAGGACCGGAGTGAGGCGG 132	Qy				
62 GGGTTGTACCGCGCGCGCGCTGAGGAAGCCACTGCAACACAGGACCGGAGTGAGGCGG 121	Db				
133 GGCAGCATGAAGCGCGCGAGGCCGCTCCATAGCGCAGCTCGGGACGCTCCGGCGGGGC 192	Qy				
122 GGCAGCATGAAGCGCGCGAGGCCGCTCCATAGCGCAGCTCGGGACGCTCCGGCGGGGC 181	Db				
193 CGGGGGAAGAAATGCAACATGCGCAGCAGCAANTGGAACAGACAGCTGGGTGTGAG 252	Qy				
182 CGGGGGAAGAAATGCAACATGCGCAGCAGCAANTGGAACAGACAGCTGGGTGTGAG 241	Db				
253 ATATTTGAAACTGCGGACTGTGAGGAGATATTGAAATCACAGGATCGGCCCTTAAATTTGGAG 312	Qy				
242 ATATTTGAAACTGCGGACTGTGAGGAGATATTGAAATCACAGGATCGGCCCTTAAATTTGGAG 301	Db				
313 CTTTTTATGTGAGCGGTATTCCTGGAGTCAAGCTTAAAGCTGCTTTGCCGATACAGA 372	Qy				
302 CTTTTTATGTGAGCGGTATTCCTGGAGTCAAGCTTAAAGCTGCTTTGCCGATACAGA 361	Db				
373 AAATATCATGGCTACATGAGTGGCTAAGGCACCATGATTTTCATGTTTGTGGAAGGAAT 432	Qy				
362 AAATATCATGGCTACATGAGTGGCTAAGGCACCATGATTTTCATGTTTGTGGAAGGAAT 421	Db				
433 GATCCAGATGAGCTTATTCAGACAGAACTTATACCTTGGCCATGCTCTGGTGAGAACAGA 492	Qy				
422 GATCCAGATGAGCTTATTCAGACAGAACTTATACCTTGGCCATGCTCTGGTGAGAACAGA 481	Db				
493 GAAATAACACTGTTTTTATCTGAAATTTCCCAAACTATCAATAGAGCAGCTCTTAATG 552	Qy				

Db	1562	TCTGAATGCAAAACAGGTTTCGGTCATTTATATACAAATTCATCTATTTTAAAGGAAGC	1621
Qy	1633	AAATATPAAACGATCCAGTGGTGGGTCCTGCTCCAAGTGATTTCAAGTGTCTTATCAAA	1692
Db	1622	AAATATPAAACGATCCAGTGGTGGGTCGCTGCTCCAAGTGATTTCAAGTGTCTTATCAAA	1681
Qy	1693	GAGGAGTAGCAATTACCAGTGGTGAATGGGAAGTTCCTGGCGGGCATGGATCTAAATATC	1752
Db	1682	GAGGAGTAGCAATTACCAGTGGTGAATGGGAAGTTCCTGGCGGGCATGGATCTAAATATC	1741
Qy	1753	CAAGTTTGATGAAGTCAGAAAGCTGGTATATTTTTGAAGGCCACCAAGAGATCCCCCTTTAGAG	1812
Db	1742	CAAGTTTGATGAAGTCAGAAAGCTGGTATATTTTTGAAGGCCACCAAGAGATCCCCCTTTAGAG	1801
Qy	1813	CATCACCTGTACGTAGTCAGTTTACGTAAATCCTGGAGAGGTGACAAAGGCTGACTGACCGT	1872
Db	1802	CATCACCTGTACGTAGTCAGTTTACGTAAATCCTGGAGAGGTGACAAAGGCTGACTGACCGT	1861
Qy	1873	GGCTACTCACATTCCTGCTGCATCAGTCAGCATGTCGCTCTTTTAAAGTAAGTATAGT	1932
Db	1862	GGCTACTCACATTCCTGCTGCATCAGTCAGCATGTCGCTCTTTTAAAGTAAGTATAGT	1921
Qy	1933	AACCCAGAGAAATCCACATGTGTGTCCTTTACAGCTATCAAGTCTCAAGTCTGAAGATGACCCA	1992
Db	1922	AACCCAGAGAAATCCACATGTGTGTCCTTTACAGCTATCAAGTCTGAAGATGACCCA	1981
Qy	1993	ACTTCGAAAAAAMAGGAATTTTGGGCCACATTTTGGATTCAGCAGGTCTCTTCCCTGAC	2052
Db	1982	ACTTCGAAAAAAMAGGAATTTTGGGCCACATTTTGGATTCAGCAGGTCTCTTCCCTGAC	2041
Qy	2053	TATACTCCTCCAGAAAATTTCTCTTTTGGAAAGTACTACTGGATTTACATGTTATGGGATG	2112
Db	2042	TATACTCCTCCAGAAAATTTCTCTTTTGGAAAGTACTACTGGATTTACATGTTATGGGATG	2101
Qy	2113	CTCTACAAGCCTCATGATCTACAGCCTCGGAAAGAAATATCCTPACTGTGCTGCTTCATATAT	2172
Db	2102	CTCTACAAGCCTCATGATCTACAGCCTCGGAAAGAAATATCCTPACTGTGCTGCTTCATATAT	2161
Qy	2173	GGTGTGCTCCTCAGGTGCAGTTGGTGAATAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTG	2232
Db	2162	GGTGTGCTCCTCAGGTGCAGTTGGTGAATAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTG	2221
Qy	2233	AATACCTTAGCCTCTCTAGGTTATGTGGTTGTAGTGATAGACAAACAGGGGATCCTGTAC	2292
Db	2222	AATACCTTAGCCTCTCTAGGTTATGTGGTTGTAGTGATAGACAAACAGGGGATCCTGTAC	2281
Qy	2293	CGAGGGCTTAAATTTGAAGGCGCTTTAAATATATAAAATGGGTCAAATAGAAATTTGACAT	2352
Db	2282	CGAGGGCTTAAATTTGAAGGCGCTTTAAATATATAAAATGGGTCAAATAGAAATTTGACAT	2341
Qy	2353	CAGGTGGAAGACTCCAAATATCTAGCTCTCGATATGATTTCAATTGACTTAGATCCGTG	2412
Db	2342	CAGGTGGAAGACTCCAAATATCTAGCTCTCGATATGATTTCAATTGACTTAGATCCGTG	2401
Qy	2413	GGCATCCAGCGTGGTCCATGAGAGATGACTCTCCCTGATGGCATTAATGCAGAGGTCA	2472
Db	2402	GGCATCCAGCGTGGTCCATGAGAGATGACTCTCCCTGATGGCATTAATGCAGAGGTCA	2461
Qy	2473	GATATCTTCAGGGTTGCTATTTGCTGGGGCCCAAGTCACTCTGTGGATCTCTTATGATACA	2532
Db	2462	GATATCTTCAGGGTTGCTATTTGCTGGGGCCCAAGTCACTCTGTGGATCTCTTATGATACA	2521
Qy	2533	GGATACCGGAAACGTTATATGGGTCACTCCCTGACCAAGATGAACAGGGCTATTACTTAGGA	2592
Db	2522	GGATACCGGAAACGTTATATGGGTCACTCCCTGACCAAGATGAACAGGGCTATTACTTAGGA	2581
Qy	2593	TCTGTGGCCATGCAAGCAGAAAAAGTTCCCTCTGAAACCAAAATCGTTTACTGCTCTTACAT	2652
Db	2582	TCTGTGGCCATGCAAGCAGAAAAAGTTCCCTCTGAAACCAAAATCGTTTACTGCTCTTACAT	2641
Qy	2653	GGTTTCTCGGATGAGAAATGTCATTTTGGACATACACAGTATATTACTGAGTTTTTTAGTG	2712
Db	2642	GGTTTCTCGGATGAGAAATGTCATTTTGGACATACACAGTATATTACTGAGTTTTTTAGTG	2701

RESULT 4
US-10-170-789-37
Sequence 37, Application US/10170789
Publication NO. US20030180930A1
GENERAL INFORMATION:
APPLICANT: Rachel E. Meyers
APPLICANT: Olandt, Peter J.
APPLICANT: Kapeller-Libermann, Rosana
APPLICANT: Curtis, Rory A. J.
APPLICANT: Williamson, Mark
APPLICANT: Welch, Nadine
TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASE, PHOSPHATASE,
AND PROTEASE FAMILY MEMBERS AND USES THEREOF
FILE REFERENCE: 10448-191001
CURRENT APPLICATION NUMBER: US/10/170,789
CURRENT FILING DATE: 2002-06-13
PRIOR APPLICATION NUMBER: US 09/797,039
PRIOR FILING DATE: 2001-02-28
PRIOR APPLICATION NUMBER: PCT/US01/06525
PRIOR FILING DATE: 2001-02-28
PRIOR APPLICATION NUMBER: US 60/186,061
PRIOR FILING DATE: 2000-02-29
PRIOR APPLICATION NUMBER: US 09/882,166
PRIOR FILING DATE: 2001-06-15
PRIOR APPLICATION NUMBER: PCT/US01/19269
PRIOR FILING DATE: 2001-06-15
PRIOR APPLICATION NUMBER: US 60/212,078
PRIOR FILING DATE: 2000-06-15
PRIOR APPLICATION NUMBER: US 09/934,406
PRIOR FILING DATE: 2001-08-21
PRIOR APPLICATION NUMBER: PCT/US01/26052
PRIOR FILING DATE: 2001-08-21
PRIOR APPLICATION NUMBER: US 60/226,740
PRIOR FILING DATE: 2000-08-21
PRIOR APPLICATION NUMBER: US 09/861,801
PRIOR FILING DATE: 2001-05-21
PRIOR APPLICATION NUMBER: PCT/US01/16549
PRIOR FILING DATE: 2001-05-21
PRIOR APPLICATION NUMBER: US 60/205,508
PRIOR FILING DATE: 2000-05-19
PRIOR APPLICATION NUMBER: US 09/801,267
PRIOR FILING DATE: 2001-03-06
PRIOR APPLICATION NUMBER: PCT/US01/07138

; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: US 60/187,454
; PRIOR FILING DATE: 2000-03-07
; PRIOR APPLICATION NUMBER: US 09/829,671
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: PCT/US01/40483
; PRIOR FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: US 60/197,508
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: US 09/961,721
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: PCT/US01/29904
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: US 60/235,023
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US 10/045,367
; PRIOR FILING DATE: 2001-11-07
; PRIOR APPLICATION NUMBER: US 60/246,561
; PRIOR FILING DATE: 2000-11-07
; PRIOR APPLICATION NUMBER: US 09/801,275
; PRIOR FILING DATE: 2001-03-06
; PRIOR APPLICATION NUMBER: PCT/US01/07074
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: US 60/187,420
; PRIOR FILING DATE: 2000-03-07
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 37
; LENGTH: 3143
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (229)...(2874)
US-10-170-789-37

Query Match 94.8%; Score 2957; DB 6; Length 3143;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3107; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy	11	GCCTCCGAGGCGAAGCGCGTCTACTGCGCGCGTCTTCTTAGTGCGCGGTTCGCGCG	70
Db	26	GCCTCCGAGGCGAAGCGCGTCTACTGCGCGCGTCTTCTTAGTGCGCGGTTCGCGCG	85
Qy	71	CTGGTGTCAACCG	130
Db	86	CTGGTGTCAACCG	145
Qy	131	GGCGCAGCATGAAGCG	190
Db	146	GGCGCAGCATGAAGCG	205
Qy	191	GCCGGGGGAGGAAATGCAATGCGCAGCAGCAATGGAACAGACAGCTGGGTGTG	250
Db	206	GCCGGGGGAGGAAATGCAATGCGCAGCAGCAATGGAACAGACAGCTGGGTGTG	265
Qy	251	AGATATTTGAACTGCGGACTGTGAGGAGAAATTAATGAAATCAGGATCGGCTAAATGG	310
Db	266	AGATATTTGAACTGCGGACTGTGAGGAGAAATTAATGAAATCAGGATCGGCTAAATGG	325
Qy	311	AdCCTTTTATGTGTAGCGGTATTCCTGGAGTCAGCTTAAAGCTGCTGCCGATACCA	370
Db	326	AdCCTTTTATGTGTAGCGGTATTCCTGGAGTCAGCTTAAAGCTGCTGCCGATACCA	385
Qy	371	GAAATATCATGGCTACATGATGGCTAAGCCACCATGATTTTCATGTTGTGAAGAGGA	430
Db	386	GAAATATCATGGCTACATGATGGCTAAGCCACCATGATTTTCATGTTGTGAAGAGGA	445
Qy	431	ATGATCCAGATGGACCTCATTCAGACAGAAATCTATTACCTTGCATGTCTGGTGAACA	490
Db	446	ATGATCCAGATGGACCTCATTCAGACAGAAATCTATTACCTTGCATGTCTGGTGAACA	505
Qy	491	GAGAAATACACTGTTTATCTGAAATTTCCCAAACTATCAATAGACGACGCTTAA	550

Db	506	GAGAAATACACTGTTTATCTGAAATTTCCCAAACTATCAATAGACGACGCTTAA	565
Qy	551	TGCTCTCTTGGAGCGCTCTTTTGGATCTTTTTCAGGCAACACTGACATATGAAATTT	610
Db	566	TGCTCTCTTGGAGCGCTCTTTTGGATCTTTTTCAGGCAACACTGACATATGAAATTT	625
Qy	611	CTCGAGAGAGAACTATTAAAGAGAAAGAAACGCAATTTGGACACAGCTCGGAATTTGCTTCT	670
Db	626	CTCGAGAGAGAACTATTAAAGAGAAAGAAACGCAATTTGGACACAGCTCGGAATTTGCTTCT	685
Qy	671	ACGATTATCACCAAGGAAGTGGAAACATTTCTGTTTCAAGCCGGTAGTGGAAATTTATCACG	730
Db	686	ACGATTATCACCAAGGAAGTGGAAACATTTCTGTTTCAAGCCGGTAGTGGAAATTTATCACG	745
Qy	731	TAAAGATGAGGGGCCACCAAGGATTTACGCAACAACCTTTAAGGCCCAATCTAGTGGAAA	790
Db	746	TAAAGATGAGGGGCCACCAAGGATTTACGCAACAACCTTTAAGGCCCAATCTAGTGGAAA	805
Qy	791	CTAGTTGTCCCAACATACCGATGGATCCAAATTTATGCCCGCTGATCCAGACTGGATTG	850
Db	806	CTAGTTGTCCCAACATACCGATGGATCCAAATTTATGCCCGCTGATCCAGACTGGATTG	865
Qy	851	CTTTTATACATACGACGATTTTGGATATCTAAACATGTAACACAGAGAAAGAGAC	910
Db	866	CTTTTATACATACGACGATTTTGGATATCTAAACATGTAACACAGAGAAAGAGAGAC	925
Qy	911	TCAGTTATGTGCAATAGTAGTAGCCCAACATCGAAGAGATGCCAGATCAGCTGGAGTCG	970
Db	926	TCAGTTATGTGCAATAGTAGTAGCCCAACATCGAAGAGATGCCAGATCAGCTGGAGTCG	985
Qy	971	CTACCTTTTCTCCAAAGAAATTTGATAGATATTTCTGGCTATTGGTGTGCCAAAG	1030
Db	986	CTACCTTTTCTCCAAAGAAATTTGATAGATATTTCTGGCTATTGGTGTGCCAAAG	1045
Qy	1031	CTGAAACAACTCCCAAGTGGTAAATTTCTAGATATTTCTATATGAAGAAATGATGAAT	1090
Db	1046	CTGAAACAACTCCCAAGTGGTAAATTTCTAGATATTTCTATATGAAGAAATGATGAAT	1105
Qy	1091	CTGAGGTGGAATTTATTTACATTCCTTACATTCCTTATGTTGAAACAGGAGGCGAGTTT	1150
Db	1106	CTGAGGTGGAATTTATTTACATTCCTTACATTCCTTATGTTGAAACAGGAGGCGAGTTT	1165
Qy	1151	TCCGTATCTCTAAACAGGTACAGCAATCTCTAAAGTCACTTTTAAAGATGTCAGAAATA	1210
Db	1166	TCCGTATCTCTAAACAGGTACAGCAATCTCTAAAGTCACTTTTAAAGATGTCAGAAATA	1225
Qy	1211	TGATTGATGCTGAAGGAAGGATCATAGATGCTCATAGATAAGGAACCTAATTTCAACCTTTG	1270
Db	1226	TGATTGATGCTGAAGGAAGGATCATAGATGCTCATAGATAAGGAACCTAATTTCAACCTTTG	1285
Qy	1271	AGATTCTATTGGAAGGATTTGAAATATATTTCCAGAGCTGGATGGACTCCTGAGGGAAT	1330
Db	1286	AGATTCTATTGGAAGGATTTGAAATATATTTCCAGAGCTGGATGGACTCCTGAGGGAAT	1345
Qy	1331	ATGTTGGTCCATCTCTACATGATCGCTCCAGACTCGCCTACAGATAGTGTGATCTCAC	1390
Db	1346	ATGTTGGTCCATCTCTACATGATCGCTCCAGACTCGCCTACAGATAGTGTGATCTCAC	1405
Qy	1391	CTGAAATTTATTTATCCGATGAGAAGATGTTTATGGAAGGCGAGACTCATTTGAGTCAG	1450
Db	1406	CTGAAATTTATTTATCCGATGAGAAGATGTTTATGGAAGGCGAGACTCATTTGAGTCAG	1465
Qy	1451	TGCTCTGATTCTGTGACGCCCACTAATTTATCTATGAGAAACACACACATCTGGATAAATA	1510
Db	1466	TGCTCTGATTCTGTGACGCCCACTAATTTATCTATGAGAAACACACACATCTGGATAAATA	1525
Qy	1511	TCCATGACATCTTTCATGTTTTTTTCCCAAGTCCAGAGGAAATTTGAGTTTATTTTGG	1570
Db	1526	TCCATGACATCTTTCATGTTTTTTTCCCAAGTCCAGAGGAAATTTGAGTTTATTTTGG	1585
Qy	1571	CCTCTGAAATGCAAAACAGGTTTCCGTCATTTATACAAAATTCATCTATTTTAAAGGAAA	1630

Db 1586 CCTGTAATGCAAAACAGGTTTCGGTCAITTTATACAAAATTACATCTATTTTAAAGGAAA 1645
Qy 1631 GCRAATATAACGATCCAGTGGTGGCTGCTGCTCCAAAGTGAATTTCAAGTGTCTATCA 1690
Db 1646 GCAATATAAACGATCCAGTGGTGGCTGCTGCTCCAAAGTGAATTTCAAGTGTCTATCA 1705
Qy 1691 AAGAGGAGATAGCAATTTACAGTGGTGAATGGGAAGTCTTGGCCGGCATGGATCTAATA 1750
Db 1706 AAGAGGAGATAGCAATTTACAGTGGTGAATGGGAAGTCTTGGCCGGCATGGATCTAATA 1765
Qy 1751 TCCAAGTTGATGAAGTCAGAAGGCTGGTATATTTTGAAGGCCAACAAAGACTCCCTTTAG 1810
Db 1766 TCCAAGTTGATGAAGTCAGAAGGCTGGTATATTTTGAAGGCCAACAAAGACTCCCTTTAG 1825
Qy 1811 AGCATCACTGTACGTAGTCAGTTACGTAATCTCTGGAGAGTGACAAAGGCTGACTGACC 1870
Db 1826 AGCATCACTGTACGTAGTCAGTTACGTAATCTCTGGAGAGTGACAAAGGCTGACTGACC 1885
Qy 1871 GTGGCTACTCACATTTCTTGCTGCATCAGTCAGCACTGTGACTTCTTTATAGTAAGTATA 1930
Db 1886 GTGGCTACTCACATTTCTTGCTGCATCAGTCAGCACTGTGACTTCTTTATAGTAAGTATA 1945
Qy 1931 GTAAACAGAGAATCCACACTGTGTCTCCCTTTTACAAGCTATCAAGTCTCTGAAGTACACC 1990
Db 1946 GTAAACAGAGAATCCACACTGTGTCTCCCTTTTACAAGCTATCAAGTCTCTGAAGTACACC 2005
Qy 1991 CAACCTTCAAAAACAAAGGATTTTGGGCCACCAATTTGGATTCAGCAGGTCCTCTCCCTG 2050
Db 2006 CAACCTTCAAAAACAAAGGATTTTGGGCCACCAATTTGGATTCAGCAGGTCCTCTCCCTG 2065
Qy 2051 ACTATACCTCTCCAGAAATTTCTCTTTGGAAGTACTACTGGATTTACATGTATGGGA 2110
Db 2066 ACTATACCTCTCCAGAAATTTCTCTTTGGAAGTACTACTGGATTTACATGTATGGGA 2125
Qy 2111 TGCTCTACAAGCCTCATGATCTACAGCCTCGGAAAGAAATATCTTACTGTGCTGTTCATAT 2170
Db 2126 TGCTCTACAAGCCTCATGATCTACAGCCTCGGAAAGAAATATCTTACTGTGCTGTTCATAT 2185
Qy 2171 ATGCTGTCTCAGGTGCAAGTGTGGTAATAATCGGTTTAAAGGAGTCAAGTATTTCCGCT 2230
Db 2186 ATGCTGTCTCAGGTGCAAGTGTGGTAATAATCGGTTTAAAGGAGTCAAGTATTTCCGCT 2245
Qy 2231 TGAATACCCCTAGCCTCTAGGTTATGTGGTTGTAGTGATAGACAACAGGGGATCTCTGTC 2290
Db 2246 TGAATACCCCTAGCCTCTAGGTTATGTGGTTGTAGTGATAGACAACAGGGGATCTCTGTC 2305
Qy 2291 ACCGAGGCTTAAATTTGAAGCGCCTTTAAATATAAAATGGGTCAAATAGAAATGACG 2350
Db 2306 ACCGAGGCTTAAATTTGAAGCGCCTTTAAATATAAAATGGGTCAAATAGAAATGACG 2365
Qy 2351 ATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTTCATTGACTTAGATCGTG 2410
Db 2366 ATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTTCATTGACTTAGATCGTG 2425
Qy 2411 TGGGCATCCACGGCTGGTCTCTATGAGGATACCTCTCCCTGATGGCATTATGCAGAGT 2470
Db 2426 TGGGCATCCACGGCTGGTCTCTATGAGGATACCTCTCCCTGATGGCATTATGCAGAGT 2485
Qy 2471 CAGATATCTTCAGGTTGCTATTTCTGGGCCCCAGTCATCTGTGGATCTTCTATGATA 2530
Db 2486 CAGATATCTTCAGGTTGCTATTTCTGGGCCCCAGTCATCTGTGGATCTTCTATGATA 2545
Qy 2531 CAGATACACGNAAGCTTATATGGGTCAACCTGACAGATGAACAGGGCTATTACTTAG 2590
Db 2546 CAGATACACGNAAGCTTATATGGGTCAACCTGACAGATGAACAGGGCTATTACTTAG 2605
Qy 2591 GATCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTGAAACCAATTCGTTTACTGCTTTAC 2650
Db 2606 GATCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTGAAACCAATTCGTTTACTGCTTTAC 2665
Qy 2651 ATGGTTTCTTGGATAGAAATGTCATTTTGCACATACAGPATATTAATCAGTATTTTATAG 2710
Db 2666 ATGGTTTCTTGGATAGAAATGTCATTTTGCACATACAGPATATTAATCAGTATTTTATAG 2725

RESULT 5

US-09-976-674-12
; Sequence 12, Application US/09976674
; Patent No. US20020115843A1
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 4829
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-976-674-12

Query Match 92.6%; Score 2889; DB 3; Length 4829;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3119; Conservative 0; Mismatches 1; Indels 2; Gaps 1;

Qy 1 AAGTGCTAAAGCCTCCGAGGCCAAGCGCGTGTCTACTGCGCCCGCTGCTTCTTAGTGCCG 60
Db 1 AAGTGCTAAAGCCTCCGAGGCCAAGCGCGTGTCTACTGCGCCCGCTGCTTCTTAGTGCCG 60
Qy 61 CGTTTCGCGCGCTGGGTGTGTCACCGCGCGCGCGAGGAAGCCACTGCAACCAAGGACCG 120
Db 61 CGTTTCGCGCGCGTGGGTGTGTCACCGCGCGCGCGAGGAAGCCACTGCAACCAAGGACCG 120
Qy 121 GAGTGAGCGCGCGCAGCATGAAGCGCGCGCGCGCGCTCCATAGCGCAGCTCGGGACGG 180
Db 121 GAGTGAGCGCGCGCAGCATGAAGCGCGCGCGCGCGCTCCATAGCGCAGCTCGGGACGG 180
Qy 181 TCCGGCGCGCGCGGGGGGAGGAAAATGCAACATGGCAGCAGCAATGGAAAACAGAACAG 240
Db 181 TCCGGCGCGCGCGGGGGGAGGAAAATGCAACATGGCAGCAGCAATGGAAAACAGAACAG 240

Qy	241	CTGGGTGTTGAGATATTTGAAACTGCGGACGTGTGAGGAGAAATATTGAATCAACAGATCGG	300
Db	241		
Qy	301	CCTAAATTTGGAGCCTTTTATGTTTGAGCGGTATTTCTGGAGTCAGCTTAAAAAGCTGCTT	360
Db	301		
Qy	361	GCGGATACCCAGAAAAATATCATGCGCTACATGCGCTAAGSCACCAATGATTTTCATGTTT	420
Db	361		
Qy	421	GTGAAGAGGAATGATCCAGATGGACCTCATTCAGACAGAACTATTATACCTTGCCTGCTCT	480
Db	421		
Qy	481	GGTGAAACAGAGAAAAATPACACTGTTTTATTCTGAAATTTCCAAAACTATCAATAGAGCA	540
Db	481		
Qy	541	GCAGTCTTAAATGCTCTCTTTGGAAGCCTCTTTTGGATCTTTTTCAGGCAACACTGGACTAT	600
Db	541		
Qy	601	GGATGCTATTCTCGAGAGAGAACTATTAAAGRAAGAAACGATTTGGACACTCGGA	660
Db	601		
Qy	661	ATTGCTTCTACGATTATCACCAAGGAAGTGGAAACATTTCTGTTTCAAGCCGGTAGTGG	720
Db	661		
Qy	721	ATTATACGTAAAGATGGAGGCCACAAGAGATTTACGCAACAACCTTTAAGGCCCAAT	780
Db	721		
Qy	781	CTAGTGGAAACTAGTTGTGCCAATACGGATGGATCCAAAAATTATGCCCTGCTGATCCA	840
Db	781		
Qy	841	GACTGGATTGCTTTTATACATAGCAACGATTTTGGATATCTAAACATCTGTAACAGAGAA	900
Db	841		
Qy	901	GAAAGGAGACTCACTTATGTGCCAATGAGCTAGCCAAATGGAAGAGAGATGCCAGATCA	960
Db	901		
Qy	961	GCTGGAGTCGCTACCTTTGTTCTCCAAGAGAAATTTGATAGATATTTCTGGCTATTGTGG	1020
Db	961		
Qy	1021	TGTCCAAAAGCTGAAACAACTCCAGTGGTGGTAAAAATTTCTAGAAATTTCTATATCAAGAA	1080
Db	1021		
Qy	1081	AATGATGAATCTGAGGTGAAAATTTATTCTCATCCCTCATGTTGAAACCAAGGAGG	1140
Db	1081		
Qy	1141	GCAGATTCAATCCGTTATCTTAAACAGGTACAGCAAAATCCTTAAAGTCACTTTTAAAGATG	1200
Db	1141		
Qy	1201	TGAGAAATTAAGATTGATGCTGAAGGAGGATCATAGATGTCATAGATAAGGAACTAATT	1260
Db	1201		
Qy	1261	CAACCTTTTGAGATTTCTATTTTGAAGGAGTTTGAATATATTTGCCAGAGCTGGATGGACTCT	1320
Db	1261		
Qy	1321	GAGGGAAAAATATGCTTGGTCCATCTACTAGATCGCTCCAGACTCGGCTACAGATAGTG	1380

[illegible]

Db 961 GCTGGAGTCGTACTTTTGTCTCCAAAGAAATTTGATAGATATCTGGCTATTGGTGG 1020
Qy
1021 TGTCCAAAAGCTGAACAACTCCAGTGGTGTGTAATTTCTAGAAATTCATATGAAGAA 1080
Db
1021 TGTCCAAAAGCTGAACAACTCCAGTGGTGTGTAATTTCTAGAAATTCATATGAAGAA 1080
Qy
1081 AATGATCAATCTCAGGTGGAAATTTATTCATGTTTACATCCCTATGTTGGTGGAAACAAGGAGG 1140
Db
1081 AATGATCAATCTCAGGTGGAAATTTATTCATGTTTACATCCCTATGTTGGTGGAAACAAGGAGG 1140
Qy
1141 GCAGATTCATTTCCGTTATCTCTAAACACAGGTACAGCAATCTCTAAAGTCACATTTTAAAGATG 1200
Db
1141 GCAGATTCATTTCCGTTATCTCTAAACACAGGTACAGCAATCTCTAAAGTCACATTTTAAAGATG 1200
Qy
1201 TCAGAAATAATGATGATGCTGAAGGAAGGATCATAGATGTCATAGATAGGAATCTAATTT 1260
Db
1201 TCAGAAATAATGATGATGCTGAAGGAAGGATCATAGATGTCATAGATAGGAATCTAATTT 1260
Qy
1261 CAACCTTTTGAAGATCTATTTGAAGGATTTGAATATATTTGCCAGAGCTGGATGACCTCCT 1320
Db
1261 CAACCTTTTGAAGATCTATTTGAAGGATTTGAATATATTTGCCAGAGCTGGATGACCTCCT 1320
Qy
1321 GAGGGAATAATGCTGGTCCATCTCTACTAGATCGCTCCAGACTCGCCTACAGATAGTG 1380
Db
1321 GAGGGAATAATGCTGGTCCATCTCTACTAGATCGCTCCAGACTCGCCTACAGATAGTG 1380
Qy
1381 TTGATCTCACTGAAATTTATTTATCCAGTAGAAGATGATGTTATGGAAGGACAGACTC 1440
Db
1381 TTGATCTCACTGAAATTTATTTATCCAGTAGAAGATGATGTTATGGAAGGACAGACTC 1440
Qy
1441 ATTGAGTCAGTGCCTGATCTGTGAGCCCACTAATTTATCTATGAAGAAACAACAGACATC 1500
Db
1441 ATTGAGTCAGTGCCTGATCTGTGAGCCCACTAATTTATCTATGAAGAAACAACAGACATC 1500
Qy
1501 TGGATAAATATCCATGACATCTTTTCATGTTTTTCCCAAGTCACGAAGAGGAATTTGAG 1560
Db
1501 TGGATAAATATCCATGACATCTTTTCATGTTTTTCCCAAGTCACGAAGAGGAATTTGAG 1560
Qy
1561 TTTATTTTTTGGCTCTGAAATGCAAAACAGGTTTTCCGTCATTTATACAAATTTACATCTATT 1620
Db
1561 TTTATTTTTTGGCTCTGAAATGCAAAACAGGTTTTCCGTCATTTATACAAATTTACATCTATT 1620
Qy
1621 TTTAAAGGAAGCAAAATATAAACCAGTCCAGTGGTGGCTCCCTGCTCCCAAGTGATTTCAAG 1680
Db
1621 TTTAAAGGAAGCAAAATATAAACCAGTCCAGTGGTGGCTCCCTGCTCCCAAGTGATTTCAAG 1680
Qy
1681 TGTCTCTATCAAGAGGAGATAGCAATTTACAGTGGTGAATGGGAAGTTCTTGGCGGCGAT 1740
Db
1681 TGTCTCTATCAAGAGGAGATAGCAATTTACAGTGGTGAATGGGAAGTTCTTGGCGGCGAT 1740
Qy
1741 GATCTAATATPCCAAGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGGCCACCAAGAC 1800
Db
1741 GATCTAATATPCCAAGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGGCCACCAAGAC 1800
Qy
1801 TCCCTTTTAGAGCATCACTGTACGTAGTCAGTTAGTAAATCTCTGGAGAGGTGACAAGG 1860
Db
1801 TCCCTTTTAGAGCATCACTGTACGTAGTCAGTTAGTAAATCTCTGGAGAGGTGACAAGG 1860
Qy
1861 CTGATGACCGTGGCTACTCACAATTTCTGCTGATCAGTCAGCATGTGACTCTTTTATA 1920
Db
1861 CTGATGACCGTGGCTACTCACAATTTCTGCTGATCAGTCAGCATGTGACTCTTTTATA 1920
Qy
1921 AATAGATATAGTAAACAGAGAAATTTGGGCCACCAATTTTGGATTCAGCAGGT 1980
Db
1921 AATAGATATAGTAAACAGAGAAATTTGGGCCACCAATTTTGGATTCAGCAGGT 1980
Qy
1981 GAAGATGACCCCACTCTCAAAAACAAGGAATTTGGGCCACCAATTTTGGATTCAGCAGGT 2040
Db
1981 GAAGATGACCCCACTCTCAAAAACAAGGAATTTGGGCCACCAATTTTGGATTCAGCAGGT 2040
Qy
2041 CCTCTTCTGACTACTCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGATTTACA 2100
Db
2041 CCTCTTCTGACTACTCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGATTTACA 2100

Qy 2101 TTGTATGGGATGCTCTACAGCCTCATGATCTACAGCCTGGAAAGAAATATCCTACTGTG 2160
Db
2101 TTGTATGGGATGCTCTACAGCCTCATGATCTACAGCCTGGAAAGAAATATCCTACTGTG 2160
Qy 2161 CTGTTTCATATATATGGTGG--TCCTCAGGTGCTGGTGAATTAATCGGTTTAAAGGAGTCA 2218
Db
2161 CTGTTTCATATATATGGTGGTCTCTCAGGTGCTGGTGAATTAATCGGTTTAAAGGAGTCA 2220
Qy 2219 AGTATTTCCGTTGAATPACCTAGCCTCTCTAGGTATGTTGTTAGTGTAGTAGACAACA 2278
Db
2221 AGTATTTCCGTTGAATPACCTAGCCTCTCTAGGTATGTTGTTAGTGTAGTAGACAACA 2280
Qy 2279 GGGGATCCTCTCACCGAGGGCTTTAAATTTGAAGGGCCCTTTTAAATATATAATGGGTCAAA 2338
Db
2281 GGGGATCCTCTGTACCGAGGGCTTTAAATTTGAAGGGCCCTTTTAAATATATAATGGGTCAAA 2340
Qy 2339 TAGAAATTTGACGATCAGGTGGAAGGACTCCAATATCTAGCTTTCTCGATATGATTTCAATG 2398
Db
2341 TAGAAATTTGACGATCAGGTGGAAGGACTCCAATATCTAGCTTTCTCGATATGATTTCAATG 2400
Qy 2399 ACTTAGATCTGTGGGCTCCACGGCTGGTCTCTATGGAGGATPACCTCTCCCTGATGGCAT 2458
Db
2401 ACTTAGATCTGTGGGCTCCACGGCTGGTCTCTATGGAGGATACCTCTCCCTGATGGCAT 2460
Qy 2459 TAATGACAGAGTCCAGATATCTTCAGGTTGCTATTTGCTGGGGCCCGCAGTCACTCTGTGGA 2518
Db
2461 TAATGACAGAGTCCAGATATCTTCAGGTTGCTATTTGCTGGGGCCCGCAGTCACTCTGTGGA 2520
Qy 2519 TCTTCTATGATACAGGATACACGGAACGTTATATGTTGTTACCTGACCAAGATGACAGG 2578
Db
2521 TCTTCTATGATACAGGATACACGGAACGTTATATGTTGTTACCTGACCAAGATGACAGG 2580
Qy 2579 GCTATTTACTTAGGATCTGTGGCCATGCAAGCAAGAAAGTTCCCTCTGCAACCAATCGTT 2638
Db
2581 GCTATTTACTTAGGATCTGTGGCCATGCAAGCAAGAAAGTTCCCTCTGCAACCAATCGTT 2640
Qy 2639 TACTGCTCTTACATGTTTCTCGATGAAGTCAATTTTGCACATACAGTATATTAC 2698
Db
2641 TACTGCTCTTACATGTTTCTCGATGAAGTCAATTTTGCACATACAGTATATTAC 2700
Qy 2699 TGAGTTTTTTAGTGAGGCTGGAAAGCCATGATTTTACAGATCTATCTCTCAGGAGAGAC 2758
Db
2701 TGAGTTTTTTAGTGAGGCTGGAAAGCCATGATTTTACAGATCTATCTCTCAGGAGAGAC 2760
Qy 2759 ACAGATAAGAGTTCTGAAATCGGGAGAACATTTGAACTGCACTCTTTTGCACTACCTTC 2818
Db
2761 ACAGATAAGAGTTCTGAAATCGGGAGAACATTTGAACTGCACTCTTTTGCACTACCTTC 2820
Qy 2819 AAGAAACCTTGGATCAGTATTTGCTGCTTAAAGTGATATAATTTGACCTGTGTAGA 2878
Db
2821 AAGAAACCTTGGATCAGTATTTGCTGCTTAAAGTGATATAATTTGACCTGTGTAGA 2880
Qy 2879 ACTCTCTGGTATACACTGGCTATTTAAACCAATAGAGGAGTTTAAATCAACAGAAACACA 2938
Db
2881 ACTCTCTGGTATACACTGGCTATTTAAACCAATAGAGGAGTTTAAATCAACAGAAACACA 2940
Qy 2939 GAAATGATCATCAATTTTGATACCTGCCATGTGTAACATCTACTCTCTGAAATATAATGTGG 2998
Db
2941 GAAATGATCATCAATTTTGATACCTGCCATGTGTAACATCTACTCTGAAATATAATGTGG 3000
Qy 2999 TGCCATGACGGGTCTACGGTTTTGTGGTATTAATCTTAATCTTAACCTTAACCCCAATCTCAA 3058
Db
3001 TGCCATGACGGGTCTACGGTTTTGTGGTATTAATCTTAATCTTAACCTTAACCCCAATCTCAA 3060
Qy 3059 AATCAATGATACATATTTCTGAGAGACCCAGCAATACCATTAAGATTTACTTAAAAAANA 3118
Db
3061 AATCAATGATACATATTTCTGAGAGACCCAGCAATACCATTAAGATTTACTTAAAAAANA 3120
Qy 3119 AA 3120
Db 3121 AA 3122

Db	961	GCATAATCTTAAAGTACCTTTTAAAGATGTCAGAAATAATGATGCTGAAGAGGATC	1020
Qy	1234	ATAGATGTCATAGATAAGAACTAATTAACCTTTTGAAGATTTCTATTTGAAGAGTTGAA	1293
Db	1021	ATAGATGTCATAGATAAGAACTAATTAACCTTTTGAAGATTTCTATTTGAAGAGTTGAA	1080
Qy	1294	TATATTTGCCAGAGCTGGATGGATCTCTGAGGGAAATATGCTTGGTCCATCTCTATAGAT	1353
Db	1081	TATATTTGCCAGAGCTGGATGGATCTCTGAGGGAAATATGCTTGGTCCATCTCTATAGAT	1140
Qy	1354	CGCTCCAGACTCGCTGAGAGTGTGATCTCACTGAAATATTTATCCAGTAGAA	1413
Db	1141	CGCTCCAGACTCGCTGAGAGTGTGATCTCACTGAAATATTTATCCAGTAGAA	1200
Qy	1414	GATGATGTTATGGAAGGCAGAGCTCACTGAGTCAGTCTGATTTCTGTGAGCCACTA	1473
Db	1201	GATGATGTTATGGAAGGCAGAGCTCACTGAGTCAGTCTGATTTCTGTGAGCCACTA	1260
Qy	1474	ATTATCTATGAAGAAACACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTTT	1533
Db	1261	ATTATCTATGAAGAAACACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTTT	1320
Qy	1534	CCCAAAGTCACGAAGAGAAATTCAGTTTATTTTGGCTCTGAATGCAAAACAGTTTC	1593
Db	1321	CCCAAAGTCACGAAGAGAAATTCAGTTTATTTTGGCTCTGAATGCAAAACAGTTTC	1380
Qy	1594	CGTCATTTATACAAATTTACATCTATTTTAAAGGAAGCAATATAACGATCCAGTGT	1653
Db	1381	CGTCATTTATACAAATTTACATCTATTTTAAAGGAAGCAATATAACGATCCAGTGT	1440
Qy	1654	GGCTGCTGCTCCAGTCAATTTCAAGTGTCTATCAAGAGGAGATAGCAATTTACAGT	1713
Db	1441	GGCTGCTGCTCCAGTCAATTTCAAGTGTCTATCAAGAGGAGATAGCAATTTACAGT	1500
Qy	1714	GATGAATGGGAAGTCTTGGCCGCAATGATCAATATCCAAAGTGTGATGAAGTCAGAAG	1773
Db	1501	GATGAATGGGAAGTCTTGGCCGCAATGATCAATATCCAAAGTGTGATGAAGTCAGAAG	1560
Qy	1774	CTGGTATATTTGAAGGCCCAAGATCCCTTTAGAGCATCACTGTGATGATCAGT	1833
Db	1561	CTGGTATATTTGAAGGCCCAAGATCCCTTTAGAGCATCACTGTGATGATCAGT	1620
Qy	1834	TACGTAAATCTGGAGAGTGACAGGCTGACCTGAGTACCTGCTACTCACATTTCTGTCG	1893
Db	1621	TACGTAAATCTGGAGAGTGACAGGCTGACCTGAGTACCTGCTACTCACATTTCTGTCG	1680
Qy	1894	ATCAGTCAGCACTGTGACTCTTTTAAAGTAAGTATAGTAACCAAGAAATCCACACTGT	1953
Db	1681	ATCAGTCAGCACTGTGACTCTTTTAAAGTAAGTATAGTAACCAAGAAATCCACACTGT	1740
Qy	1954	GTGTCCCTTTTACAAAGCTATCAAGTCTCTGAAGATGACCCAACTTGCAAAACAAAGAAATTT	2013
Db	1741	GTGTCCCTTTTACAAAGCTATCAAGTCTCTGAAGATGACCCAACTTGCAAAACAAAGAAATTT	1800
Qy	2014	TGGGCCACCAATTTTGGATTCAGAGTCTCTCTCTGATATATCTCTCCAGAAATTTTC	2073
Db	1801	TGGGCCACCAATTTTGGATTCAGAGTCTCTCTCTGATATATCTCTCCAGAAATTTTC	1860
Qy	2074	TCCTTTGAAAGTACTACTGATTTACATTTGATGGATGCTCTACAGCCCTCATGATCTA	2133
Db	1861	TCCTTTGAAAGTACTACTGATTTACATTTGATGGATGCTCTACAGCCCTCATGATCTA	1920
Qy	2134	CAGCTGGAAGAAATCTACTGTGCTGCTCATATATGTTGGTCCCTCAGTGCAGTTG	2193
Db	1921	CAGCTGGAAGAAATCTACTGTGCTGCTCATATATGTTGGTCCCTCAGTGCAGTTG	1980
Qy	2194	GTGAATATCGTTTTAAAGAGTCAAGTATTTCCGCTTGAATACCCCTAGCCCTCTAGGT	2253
Db	1981	GTGAATATCGTTTTAAAGAGTCAAGTATTTCCGCTTGAATACCCCTAGCCCTCTAGGT	2040
Qy	2254	TATGTGTTGTAGTGATAGACAACAGGGATCTGTACCCGAGGCTTAAATTTGAAGGC	2313
Db	2314	GCCTTTAATATAAAATGGGTCAAAATGAATGAACATGACATCAGGTGGAAGGACTCCCAATAT	2373
Qy	2101	GCCTTTAATATAAAATGGGTCAAAATGAATGAACATGACATCAGGTGGAAGGACTCCCAATAT	2160
Db	2374	CTAGCTTCTCGATATGATTTTCAATGACTTAGATCGTGTGGGCATCCACGGCTGGTCTAT	2433
Qy	2161	CTAGCTTCTCGATATGATTTTCAATGACTTAGATCGTGTGGGCATCCACGGCTGGTCTAT	2220
Db	2434	GGAGGATACCTCTCCCTGATGGCAATTAATGAGAGTCAAGATATCTTCAGGGTTGCTATT	2493
Qy	2221	GGAGGATACCTCTCCCTGATGGCAATTAATGAGAGTCAAGATATCTTCAGGGTTGCTATT	2280
Db	2494	GCTGGGGCCCCAGTCACTCTGTGGATCTTTATGATACAGGATACACGGAAACGTTATATG	2553
Qy	2281	GCTGGGGCCCCAGTCACTCTGTGGATCTTTATGATACAGGATACACGGAAACGTTATATG	2340
Db	2554	GCTCACCTGACCAAGTAACAGGGCTATTACTTAGATCTGTGGCATGCAAGCAGAA	2613
Qy	2341	GCTCACCTGACCAAGTAACAGGGCTATTACTTAGATCTGTGGCATGCAAGCAGAA	2400
Db	2614	AGTTCCCTCTGAAACCAAAATCGTTTACTCTTACATGGTTTCTGGATGAGATGTC	2673
Qy	2401	AGTTCCCTCTGAAACCAAAATCGTTTACTCTTACATGGTTTCTGGATGAGATGTC	2460
Db	2674	CATTTTGCACATACACAGTATATTTACTGAGTTTTTTTAGTGGGCTGGAAGCCATATGAT	2733
Qy	2461	CATTTTGCACATACACAGTATATTTACTGAGTTTTTTTAGTGGGCTGGAAGCCATATGAT	2520
Db	2734	TTACAGATCTATCTCCAGGAGACACAGCATAGAGTTTCTGAAATCGGAGAACATTTAT	2793
Qy	2521	TTACAGATCTATCTCCAGGAGACACAGCATAGAGTTTCTGAAATCGGAGAACATTTAT	2580
Db	2794	GAATCGACTTTTGGCACTACCTTCAAGAAAACCTTTGGATCACGTTTGTCTCTAAAA	2853
Qy	2581	GAATCGACTTTTGGCACTACCTTCAAGAAAACCTTTGGATCACGTTTGTCTCTAAAA	2640
Db	2854	GTGATATAA 2862	
Qy	2641	GTGATATAA 2649	
Db			
RESULT 8			
US-10-054-776-1			
; Sequence 1, Application US/10054776			
; Publication No. US20030165818A1			
; GENERAL INFORMATION:			
; APPLICANT: Mark Robert Edbrooke			
; APPLICANT: Alan Peter Lewis			
; TITLE OF INVENTION: NOVEL PROTEIN			
; FILE REFERENCE: Q61042US			
; CURRENT APPLICATION NUMBER: US/10/054,776			
; CURRENT FILING DATE: 2002-01-23			
; NUMBER OF SEQ ID NOS: 2			
; SOFTWARE: PatentIn version 3.0			
; SEQ ID NO 1			
; LENGTH: 2649			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
; FEATURE:			
; NAME/KEY: CDS			
; LOCATION: (1)...(2649)			
US-10-054-776-1			
Query Match			
Best Local Similarity 80.0%; Score 2496; DB 6; Length 2649;			
Matches 2646; Conservative 0; Mismatches 3; Indels 0; Gaps 0;			
Qy	214	ATGGCAGCAGCAATGGAAACAGAACAGCTGGGTGTTGAGATATTTGAAACTCGGACTGT	273
Db	1	ATGGCAGCAGCAATGGAAACAGAACAGCTGGGTGTTGAGATATTTGAAACTCGGACTGT	60

Qy	274	GAGGAGAAATATTTGAATCACAGGATCGGCCATAAATTTGGAGCCCTTTTATATGTTGAGCGGTAT	333
Db	61	GAGGAGAAATATTTGAATCACAGGATCGGCCATAAATTTGGAGCCCTTTTATATGTTGAGCGGTAT	120
Qy	334	TCCTGGAGTCAGCTTAAAAAGCTGCTTGCAGATACACAGAAAATATCATGGCTACATGATG	393
Db	121	TCCTGGAGTCAGCTTAAAAAGCTGCTTGCAGATACACAGAAAATATCATGGCTACATGATG	180
Qy	394	GCTAAGGCACCATCATGATTTTCATGTTGTGAAGAGGAATGATCCAGATGACCTCATTTCA	453
Db	181	GCTAAGGCACCATCATGATTTTCATGTTGTGAAGAGGAATGATCCAGATGACCTCATTTCA	240
Qy	454	GACAGAACTATTACCTTGCCATGTCGTGTGAGACAGAGAAAATACACGTGTTTATTCCT	513
Db	241	GACAGAACTATTACCTTGCCATGTCGTGTGAGACAGAGAAAATACACGTGTTTATTCCT	300
Qy	514	GAATTTCCCAAAACTATCAATAGACAGCAGCTTTAATGCTCTCTTGGAGCCCTCTTTTG	573
Db	301	GAATTTCCCAAAACTATCAATAGACAGCAGCTTTAATGCTCTCTTGGAGCCCTCTTTTG	360
Qy	574	GATCTTTTTTCAGGCAACCTGACATCGGAATGATTCTCGAGNAGAGAACTATTAAAGA	633
Db	361	GATCTTTTTTCAGGCAACCTGACATCGGAATGATTCTCGAGNAGAGAACTATTAAAGA	420
Qy	634	GAAGAAACCGATTTGGACACCTCGGAATTCCTTTACGATTATCACAGGAAGTGA	693
Db	421	GAAGAAACCGATTTGGACACCTCGGAATTCCTTTACGATTATCACAGGAAGTGA	480
Qy	694	ACATTTCTGTTTTCAAGCCGTTAGTGGAAATTTATCACGTAAAAAGATGGAGGGCCACAAGGA	753
Db	481	ACATTTCTGTTTTCAAGCCGTTAGTGGAAATTTATCACGTAAAAAGATGGAGGGCCACAAGGA	540
Qy	754	TTTACGCAACAACCTTTAAGGCCAATCTAGTGGAAACTAGTTGTCTCCCAACATACGGATG	813
Db	541	TTTACGCAACAACCTTTAAGGCCAATCTAGTGGAAACTAGTTGTCTCCCAACATACGGATG	600
Qy	814	GATCCAAAATTAATGCCCCGTGATCCAGACTGGATTGCTTTTATACATAGCAACGATTT	873
Db	601	GATCCAAAATTAATGCCCCGTGATCCAGACTGGATTGCTTTTATACATAGCAACGATTT	660
Qy	874	TGGATATCTAACCATCGTAAACAGAGAAAGAGAGACTCACTTATGTGCACAACTCAGCTA	933
Db	661	TGGATATCTAACCATCGTAAACAGAGAAAGAGAGACTCACTTATGTGCACAACTCAGCTA	720
Qy	934	GGCAACATGGAGAGAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCCAAGAA	993
Db	721	GGCAACATGGAGAGAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCCAAGAA	780
Qy	994	TTTGTATAGATATTTCTGGCTATTGTGTGTGTCCTCAAAAGCTGMAACAACTCCCAAGTGGT	1053
Db	781	TTTGTATAGATATTTCTGGCTATTGTGTGTGTCCTCAAAAGCTGMAACAACTCCCAAGTGGT	840
Qy	1054	AAAAATCTTAGAATTTCTATATGAGAAAATGATGAATCTGAGGTGGAAATTTATCATGTT	1113
Db	841	AAAAATCTTAGAATTTCTATATGAGAAAATGATGAATCTGAGGTGGAAATTTATCATGTT	900
Qy	1114	ACATCCCTATGTTGGRAAACAGAGGGCAGATTTCAATCCGTTATCCTTAAACAGGTACA	1173
Db	901	ACATCCCTATGTTGGRAAACAGAGGGCAGATTTCAATCCGTTATCCTTAAACAGGTACA	960
Qy	1174	GCAATCTTAAGTCACCTTTAGATGTCAGAAATTAATGATGCTGTAAGGAAGGATC	1233
Db	961	GCAATCTTAAGTCACCTTTAGATGTCAGAAATTAATGATGCTGTAAGGAAGGATC	1020
Qy	1234	ATAGTGTCTATAGATAGGAACCTAATTCACCTTTTGAGATTCCTATTTGAAGGAGTTGAA	1293
Db	1021	ATAGTGTCTATAGATAGGAACCTAATTCACCTTTTGAGATTCCTATTTGAAGGAGTTGAA	1080
Qy	1294	TATATTCGACAGCTGGATGGACTCTCTGAGGGAATAATGTGTGCTCATCTACTAGAT	1353
Db	1081	TATATTCGACAGCTGGATGGACTCTCTGAGGGAATAATGTGTGCTCATCTACTAGAT	1140
Qy	1354	CGCTCCGAGACTCGCCTACAGATAGTGTGTGATCTCACCTGGAATTTATTTATCAGTAGTAA	1413

Db	1141	CGCTCCAGACTCGCTCGAGTAGTGTGATCTCACCTGAAATATTTATTCACGAGTAGAA	1200
Qy	1414	GATGATGTTATGGAAAGCGCAGACTCATTTGAGTCAGTCGCTGATTTCTGTGAGCGCACATA	1473
Db	1201	GATGATGTTATGGAAAGCGCAGACTCATTTGAGTCAGTCGCTGATTTCTGTGAGCGCACATA	1260
Qy	1474	ATTATCTATGAAGAAAACAACAGACATCTGGATAAATATCATGACATCTTTCAATGTTTT	1533
Db	1261	ATTATCTATGAAGAAAACAACAGACATCTGGATAAATATCATGACATCTTTCAATGTTTT	1320
Qy	1534	CCCCAAAGTCAGAAAGGAAATGAGTTATTTTTGCCCTCGAAATGCAAAACAGGTTTC	1593
Db	1321	CCCCAAAGTCAGAAAGGAAATGAGTTATTTTTGCCCTCGAAATGCAAAACAGGTTTC	1380
Qy	1594	CGTCATTTATCAAAAATACATCTATTTTAAAGGAAAGCAAAATATAAACGATCCAGTGGT	1653
Db	1381	CGTCATTTATCAAAAATACATCTATTTTAAAGGAAAGCAAAATATAAACGATCCAGTGGT	1444
Qy	1654	GGCTGCTGCTCCAAGTGATTTCAAGTGTCTTATCAAAAGAGGAGATAGCAATATACCAAGT	1713
Db	1441	GGCTGCTGCTCCAAGTGATTTCAAGTGTCTTATCAAAAGAGGAGATAGCAATATACCAAGT	1500
Qy	1714	GGTAAATGGGAAGTCTTTGGCCGGCATGGATCTAAATATCAAAAGTTGATGAAGTCAGAAGG	1773
Db	1501	GGTAAATGGGAAGTCTTTGGCCGGCATGGATCTAAATATCAAAAGTTGATGAAGTCAGAAGG	1560
Qy	1774	CTGGTATATTTTGAAGGCACAAAGACTCCCCCTTTAGAGCATCACTGTACGTAGTCAGT	1833
Db	1561	CTGGTATATTTTGAAGGCACAAAGACTCCCCCTTTAGAGCATCACTGTACGTAGTCAGT	1620
Qy	1834	TACGTAATCCTGGAGAGGTGACAAAGGCTGACTGACCGTGGCTACTCACATTTCTGCTGC	1893
Db	1621	TACGTAATCCTGGAGAGGTGACAAAGGCTGACTGACCGTGGCTACTCACATTTCTGCTGC	1680
Qy	1894	ATCAGTCAGCACTGTGACTTCTTTATAAGTAAAGTAGTAGTAAGCAGCAATTCACACTGT	1953
Db	1681	ATCAGTCAGCACTGTGACTTCTTTATAAGTAAAGTAGTAGTAAGCAGCAATTCACACTGT	1740
Qy	1954	GTGTCCCTTTTACAAGCTATCAAGTCTCTGAAGATGACCCAACTTCGCAAAACAAGGAATTT	2013
Db	1741	GTGTCCCTTTTACAAGCTATCAAGTCTCTGAAGATGACCCAACTTCGCAAAACAAGGAATTT	1800
Qy	2014	TGGGCCACCAATTTTGATTTACAGAGTCTCTCTCTGACTATCTCTCCAGAAATTTTC	2073
Db	1801	TGGGCCACCAATTTTGATTTACAGAGTCTCTCTCTGACTATCTCTCCAGAAATTTTC	1860
Qy	2074	TCCTTTTGAAGTACTCTGATTTACATTTGATGGGATGCTCTACAGCCTCATGATCTA	2133
Db	1861	TCCTTTTGAAGTACTCTGATTTACATTTGATGGGATGCTCTACAGCCTCATGATCTA	1920
Qy	2134	CAGCCTGGAAAGAAATATCCTACTGTCTGTTTCATATATGTTGTTCTCTCAGGTCAGTTG	2193
Db	1921	CAGCCTGGAAAGAAATATCCTACTGTCTGTTTCATATATGTTGTTCTCTCAGGTCAGTTG	1980
Qy	2194	GTCAATAATTCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCCCTAGGCCTCTTAGGT	2253
Db	1981	GTCAATAATTCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCCCTAGGCCTCTTAGGT	2040
Qy	2254	TATGTGGTTGTATGATAGACAACAGGGGATCTCTGTCAACCGAGGCTTAAATTTGAAGGC	2313
Db	2041	TATGTGGTTGTATGATAGACAACAGGGGATCTCTGTCAACCGAGGCTTAAATTTGAAGGC	2100
Qy	2314	GCCTTTAAATATAAAATGGGTCAAATAGAAATTACGATCAGGTGGAAGGACCTCCAATAT	2373
Db	2101	GCCTTTAAATATAAAATGGGTCAAATAGAAATTACGATCAGGTGGAAGGACCTCCAATAT	2160
Qy	2374	CTAGCTTCTCGATATGATTTCAITTGACTTAGATCGTGTGGGCATCCACGGCTGGTCCAT	2433
Db	2161	CTAGCTTCTCGATATGATTTCAITTGACTTAGATCGTGTGGGCATCCACGGCTGGTCCAT	2220
Qy	2434	GGAGGATACCTCTCCCTGATGGCATTAATGCAGAGGTCAAGATATCTTTCAGGGTTGCTATT	2493

Db	2221	GGAGGATACCTCTCCCTGATGGCATTAATGCAGAGGTCAGATATCTTCAGGGTTGCTATT	2280
Qy	2494	GCTGGGGCCCCAGTCACTCTGTGGATCTCTATGATACAGGATACACGGAACGTTATATG	2553
Db	2281	GCTGGGGCCCCAGTCACTCTGTGGATCTCTATGATACAGGATACACGGAACGTTATATG	2340
Qy	2554	GGTCAACCTTGACCAAGTGAACAGGGCTATTACTTAGGATCTGTGGCCATCGAAGCAGAA	2613
Db	2341	GGTCACCTTGACCAAGTGAACAGGGCTATTACTTAGGATCTGTGGCCATCGAAGCAGAA	2400
Qy	2614	AGTTCCCCCTGTGAACCAAAATCGTTTACTGCTCTTACATCGTTTCTCGATGAGAATGTC	2673
Db	2401	AGTTCCCCCTGTGAACCAAAATCGTTTACTGCTCTTACATCGTTTCTCGATGAGAATGTC	2460
Qy	2674	CAITTTTGACATACCAGTATATTACTGAGTTTTTATGTCAGGGCTGGAAAGCCATATGAT	2733
Db	2461	CAITTTTGACATACCAGTATATTACTGAGTTTTTATGTCAGGGCTGGAAAGCCATATGAT	2520
Qy	2734	TTACAGATCATCTCTCAGGAGAGACACAGCATTAAGAGTTCCTGAATCGGGAGAACATTAT	2793
Db	2521	TTACAGATCATCTCTCAGGAGAGACACAGCATTAAGAGTTCCTGAATCTGGAGAACATTAT	2580
Qy	2794	GAACTGCATCTTTTGCATCTACCTTCAAGAAAACTTTGGATCAAGTATTCGTCTCTAAAA	2853
Db	2581	GAACTGCATCTTTTGCATCTACCTTCAAGAAAACTTTGGATCAAGTATTCGTCTCTAAAA	2640
Qy	2854	GTGATATAA	2862
Db	2641	GTGATATAA	2649

RESULT 9

```

US-09-976-674-2
; Sequence 2, Application US/09976674
; Patent No. US20020115843A1
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 2
; LENGTH: 2671
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-976-674-2

```

Qy	453	AGACAGAA	ATCTATT	TACTT	CGCCAT	GTCTG	TGAGAA	CAGAGAA	ATA	CAC	TGTTT	ATTC	512
Db	247	AGACAGAA	ATCTATT	TACTT	CGCCAT	GTCTG	TGAGAA	CAGAGAA	ATA	CAC	TGTTT	ATTC	306
Qy	513	TGAAAT	TCCAAA	ACTAT	CAAT	TAGAG	CAC	AGCTT	TAAT	GTCTCT	CGAAG	CCCTCT	572
Db	307	TGAAAT	TCCAAA	ACTAT	CAAT	TAGAG	CAC	AGCTT	TAAT	GTCTCT	CGAAG	CCCTCT	366
Qy	573	GGAT	CTTTTT	CAGG	CAACA	CTG	GACTAT	GGAAT	TGTA	TCTC	GAGA	GAAGAA	632
Db	367	GGAT	CTTTTT	CAGG	CAACA	CTG	GACTAT	GGAAT	TGTA	TCTC	GAGA	GAAGAA	426
Qy	633	AGAA	AAAA	CGCA	ATTC	GGAA	CAGT	CGGA	ATTC	GTCTT	CACAT	TAAT	692
Db	427	AGAA	AAAA	CGCA	ATTC	GGAA	CAGT	CGGA	ATTC	GTCTT	CACAT	TAAT	486
Qy	693	ACAT	TTTCT	GTGTTT	CAAG	CCGGT	AGT	GGAA	TTTAT	CAC	GTA	AAAGAT	752
Db	487	ACAT	TTTCT	GTGTTT	CAAG	CCGGT	AGT	GGAA	TTTAT	CAC	GTA	AAAGAT	546
Qy	753	ATT	TACG	CAAA	CAAA	CTTT	TAA	GGCC	CAAT	CT	TAGT	GGAAA	812
Db	547	ATT	TACG	CAAA	CAAA	CTTT	TAA	GGCC	CAAT	CT	TAGT	GGAAA	606
Qy	813	GGAT	CCAAAA	TTAT	TG	CCCG	CGT	GAT	CCAGA	CT	GGAT	TGCTTT	872
Db	607	GGAT	CCAAAA	TTAT	TG	CCCG	CGT	GAT	CCAGA	CT	GGAT	TGCTTT	666
Qy	873	TTG	GAAT	CT	TAA	CAAT	CGTA	AC	CAGAG	AA	GAGAG	ACT	932
Db	667	TTG	GAAT	CT	TAA	CAAT	CGTA	AC	CAGAG	AA	GAGAG	ACT	726
Qy	933	AGC	CAAC	AT	GGAA	GAGAT	GCC	GAT	CAG	CT	GGAGT	CGCT	992
Db	727	AGC	CAAC	AT	GGAA	GAGAT	GCC	GAT	CAG	CT	GGAGT	CGCT	786
Qy	993	ATT	TGAT	AGAT	TAT	CT	GGCT	TAT	TGGT	GGT	CTC	CAAA	1052
Db	787	ATT	TGAT	AGAT	TAT	CT	GGCT	TAT	TGGT	GGT	CTC	CAAA	846
Qy	1053	TAAA	AT	CTT	TAGA	AT	CT	TAT	TGA	AAAA	TGAT	TGAAT	1112
Db	847	TAAA	AT	CTT	TAGA	AT	CT	TAT	TGA	AAAA	TGAT	TGAAT	906
Qy	1113	TACAT	CCCC	AT	GTG	TG	GNAA	CAGAG	GGCAG	AGT	CAT	TCCG	1172
Db	907	TACAT	CCCC	AT	GTG	TG	GNAA	CAGAG	GGCAG	AGT	CAT	TCCG	966
Qy	1173	AGCA	AT	CCT	AAAGT	CACT	TTTT	TAAGT	GTG	CAG	AAAA	TAAT	1232
Db	967	AGCA	AT	CCT	AAAGT	CACT	TTTT	TAAGT	GTG	CAG	AAAA	TAAT	1026
Qy	1233	CAT	AGAT	GT	CA	TAG	TAAG	AACT	TAAT	TCA	ACCT	TTTT	1292
Db	1027	CAT	AGAT	GT	CA	TAG	TAAG	AACT	TAAT	TCA	ACCT	TTTT	1086
Qy	1293	ATA	TAT	TG	CCAG	AGCT	GTG	ACT	CTC	TG	AGG	AAAA	1352
Db	1087	ATA	TAT	TG	CCAG	AGCT	GTG	ACT	CTC	TG	AGG	AAAA	1146
Qy	1353	TG	CT	CCCC	AG	ACT	CG	CTT	CAC	AGT	AGT	TTAT	1412
Db	1147	TG	CT	CCCC	AG	ACT	CG	CTT	CAC	AGT	AGT	TTAT	1206
Qy	1413	AGAT	GA	TGTT	TG	AAAA	GGCAG	AG	ACT	CA	T	TGAGT	1472
Db	1207	AGAT	GA	TGTT	TG	AAAA	GGCAG	AG	ACT	CA	T	TGAGT	1266
Qy	1473	AA	TAT	CT	TAT	G	AAAA	CAAA	CAG	ACT	CT	CGAT	1532
Db	1267	AA	TAT	CT	TAT	G	AAAA	CAAA	CAG	ACT	CT	CGAT	1326

QY 1533 TCCCAAAGTCACGAAGAGGAAATGAGTTTATTTTGGCTCTGTAATGCAAAACAGGTTT 1592
Db 1327 TCCCAAAGTCACGAAGAGGAAATGAGTTTATTTTGGCTCTGTAATGCAAAACAGGTTT 1386
QY 1593 CCGTCATTATACAAAAATTACATCTATTTTAAAGGAAAGCAAAATATNAACGATCCAGTGG 1652
Db 1387 CCGTCATTATACAAAAATTACATCTATTTTAAAGGAAAGCAAAATATNAACGATCCAGTGG 1446
QY 1653 TGGGCTGCTGCTCCCAAGTCATTTCAAGTGTCTCTATCAAGAGGAGATGCAATATACAG 1712
Db 1447 TGGGCTGCTGCTCCCAAGTCATTTCAAGTGTCTCTATCAAGAGGAGATGCAATATACAG 1506
QY 1713 TGGTGAATGGGAAGTCTTGGCGGCGATCGATCTAATATCCAAAGTGTGATGAAGTCAGAAG 1772
Db 1507 TGGTGAATGGGAAGTCTTGGCGGCGATCGATCTAATATCCAAAGTGTGATGAAGTCAAGAAG 1566
QY 1773 GCTGGTATATTTGAAGGCCACAAAGACTCCCTTTAGAGCATACCTGTGACGTAGTCAG 1832
Db 1567 GCTGGTATATTTGAAGGCCACAAAGACTCCCTTTAGAGCATACCTGTGACGTAGTCAG 1626
QY 1833 TTACGTAAATCTCGAGAGGTGACAAAGGCTGACTGACCGTGGCTACTCACATTTCTTGCTG 1892
Db 1627 TTACGTAAATCTCGAGAGGTGACAAAGGCTGACTGACCGTGGCTACTCACATTTCTTGCTG 1686
QY 1893 CATCAGTCAGACTGTGACTTCTTTATAAGTAAGTATAGTAACCAAGAGAAATCCACACTG 1952
Db 1687 CATCAGTCAGACTGTGACTTCTTTATAAGTAAGTATAGTAACCAAGAGAAATCCACACTG 1746
QY 1953 TGTGTCCCTTTTCAAGCTATCAAGTCTCAAGTCAAGATGACCAACTTGCAGAAACAAAGGAAT 2012
Db 1747 TGTGTCCCTTTTCAAGCTATCAAGTCTCAAGTCAAGATGACCAACTTGCAGAAACAAAGGAAT 1806
QY 2013 TTGGGCGCACCATTTTGGATTCAGCAGGTCTCTCTGACTATCTACTCTCCAGAAATTTT 2072
Db 1807 TTGGGCGCACCATTTTGGATTCAGCAGGTCTCTCTGACTATCTACTCTCCAGAAATTTT 1866
QY 2073 CTCTTTTGAAGTACTCTGATTTACATTTGATGGATCTCTTACAGGCTCATGATCT 2132
Db 1867 CTCTTTTGAAGTACTCTGATTTTACATTTGATGGATCTCTTACAGGCTCATGATCT 1926
QY 2133 ACAGCTTGGAAAGAAATATCTACTGTGCTGTTCATATATGGTGGTCTCTCAGGTGCAAGTT 2192
Db 1927 ACAGCTTGGAAAGAAATATCTACTGTGCTGTTCATATATGGTGGTCTCTCAGGTGCAAGTT 1986
QY 2193 GGTGAATATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTAGCCTCTCTAGG 2252
Db 1987 GGTGAATATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTAGCCTCTCTAGG 2046
QY 2253 TTATGTGGTTGTAGTATAGACAACAGGGGATCCCTGTACCGAGGGCTTAAATTTGAAGG 2312
Db 2047 TTATGTGGTTGTAGTATAGACAACAGGGGATCCCTGTACCGAGGGCTTAAATTTGAAGG 2106
QY 2313 CGCCTTTAAATATAAAATGGGTCAAAATAGAAATTTGACGATCAGGTGGAAGGACTCCAATA 2372
Db 2107 CGCCTTTAAATATAAAATGGGTCAAAATAGAAATTTGACGATCAGGTGGAAGGACTCCAATA 2166
QY 2373 TCTAGCTTCTCGATATGATTTTCATTTGACTTTAGATCGTGTGGGATCCAGGCTGGTCCCTA 2432
Db 2167 TCTAGCTTCTCGATATGATTTTCATTTGACTTTAGATCGTGTGGGATCCAGGCTGGTCCCTA 2226
QY 2433 TGAGGATACCTCTCCCTGATGGCATTAAATCGAGGTCAGATATCTTCAAGGTTGCTAT 2492
Db 2227 TGAGGATACCTCTCCCTGATGGCATTAAATCGAGGTCAGATATCTTCAAGGTTGCTAT 2286
QY 2493 TGTCTGGGCGCCAGTCATCTGTGGATCTTCTATGATACAGGATACACGGAAGCTTATAT 2552
Db 2287 TGTCTGGGCGCCAGTCATCTGTGGATCTTCTATGATACAGGATACACGGAAGCTTATAT 2346
QY 2553 GGGTCACCTTGACCAAGATGAACAGGGCTATTTACTTAGGATCTGTGGCCATGCAAGCAGA 2612
Db 2347 GGGTCACCTTGACCAAGATGAACAGGGCTATTTACTTAGGATCTGTGGCCATGCAAGCAGA 2406
QY 2613 AAAGTTCCCTCTGAACCAAAATCGGTTTACTGCTCTTTACATGGTTTCTCTGGATGAGATGT 2672

Db 2407 AAAGTTCCCTCTGAACCAAAATCGTTTACTGCTCTTACATGGTTTCTCTGGATGAGAAATGT 2466
QY 2673 CCATTTTGCACATACCAAGTATATTTACTGAGTTTCTTGTAGGGCTGGAAGGCCATATGA 2732
Db 2467 CCATTTTGCACATACCAAGTATATTTACTGAGTTTCTTGTAGGGCTGGAAGGCCATATGA 2526
QY 2733 TTTACAGATCTATCTCTCAGGAGAGACACAGCATAAAGAGTTCCTGAATCGGAGAACATTA 2792
Db 2527 TTTACAGATCTATCTCTCAGGAGAGACACAGCATAAAGAGTTCCTGAATCGGAGAACATTA 2586
QY 2793 TGAATCGATCTTTTTCACATACCTTCAAGAAAAACCTTGGATCACGTATTTGCTCTTAAA 2852
Db 2527 TGAATCGATCTTTTTCACATACCTTCAAGAAAAACCTTGGATCACGTATTTGCTCTTAAA 2646
QY 2853 AGTGATAT 2860
Db 2647 AGTGATAT 2654
RESULT 10
US-10-982-512-2
; Sequence 2, Application US/10982512
; Publication No. US20050059081A1
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinesya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/10/982,512
; PRIOR FILING DATE: 2004-11-05
; PRIOR APPLICATION NUMBER: US/09/976,674
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US/60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 2671
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-982-512-2
Query Match 80.0%; Score 2495; DB 9; Length 2671;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2645; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 213 CATGGCAGCAGCAATGGAACAGAACACAGACAGCTGGGTGTTGAGATATTTGAAACTCGGACTG 272
Db 7 CATGGCAGCAGCAATGGAACAGAACACAGACAGCTGGGTGTTGAGATATTTGAAACTCGGACTG 66
QY 273 TGAGGAGATATTTGAATCACAGGATCGGCTTAAATTCGAGCCTTTTATGAGCGGTA 332
Db 67 TGAGGAGATATTTGAATCACAGGATCGGCTTAAATTCGAGCCTTTTATGAGCGGTA 126
QY 333 TTCTCGAGTCAGCTTAAAGAGCTGCTTCCGATACAGAAATATCATGGCTACATGAT 392
Db 127 TTCTCGAGTCAGCTTAAAGAGCTGCTTCCGATACAGAAATATCATGGCTACATGAT 186
QY 393 GGCCTAAGGCACCAATGATTTTCATGTTTGAAGAGGAATGATCCAGATCGACTCATTC 452
Db 187 GGCCTAAGGCACCAATGATTTTCATGTTTGAAGAGGAATGATCCAGATCGACTCATTC 246
QY 453 AGACAGAATCTATTACCTTGGCATGTCTGGTGAAGAACAGAGAAATACTGTTTATTTC 512
Db 247 AGACAGAATCTATTACCTTGGCATGTCTGGTGAAGAACAGAGAAATACTGTTTATTTC 306
QY 513 TGAATTTCCAAAATCTCAATAGAGCAGCAGTCTTAAATGCTCTCTTGAAGGCTCTTTT 572
Db 307 TGAATTTCCAAAATCTCAATAGAGCAGCAGTCTTAAATGCTCTCTTGAAGGCTCTTTT 366

QY	573	GGATCTTTTTCAGGCAACACTGGACTATGGAATGTATCTTCGAGAAGAAAGCACTATTAAAG	632
Db	367	GGATCTTTTTCAGGCAACACTGGACTATGGAATGTATCTTCGAGAAGAAAGCACTATTAAAG	426
QY	633	AdAAGAAACGCATTCGACAGTCGGAATTCCTTACGATTATCACCAGGAGTGG	692
Db	427	AdAAGAAACGCATTCGACAGTCGGAATTCCTTACGATTATCACCAGGAGTGG	486
QY	693	AAcATTCTCTGTTTCAAGCCGGTAGTGGAAATTTATCAGCTAAAGATGAGGCGCCACAAGG	752
Db	487	AAcATTCTCTGTTTCAAGCCGGTAGTGGAAATTTATCAGCTAAAGATGAGGCGCCACAAGG	546
QY	753	ATTTACGCAACAACTTTAAAGGCCAATCTAGTGGAAACTAGTGTCTCCAAACATACGGAT	812
Db	547	ATTTACGCAACAACTTTAAAGGCCAATCTAGTGGAAACTAGTGTCTCCAAACATACGGAT	606
QY	813	GGATCCAAATATATGCCCCTGATCCAGACTGGATGCTTTTATACATAGCAACGATAT	872
Db	607	GGATCCAAATATATGCCCCTGATCCAGACTGGATGCTTTTATACATAGCAACGATAT	666
QY	873	TTGGATATCTAACATCTTAACACAGAGCAAGAGGAGACTCACTTATGTGCAATGAGCT	932
Db	667	TTGGATATCTAACATCTTAACACAGAGCAAGAGGAGACTCACTTATGTGCAATGAGCT	726
QY	933	AdCCAAcATGGAAGATGCCAGATCAGCTGGAGTCGCTACTTTGTTCTCCAGAGA	992
Db	727	AGCCAAcATGGAAGATGCCAGATCAGCTGGAGTCGCTACTTTGTTCTCCAGAGA	786
QY	993	ATTTGATAGATATTTCTGGCTATTGTTGGTGTCCAAAGCTGAAACAACTCCAGTGGTGG	1052
Db	787	ATTTGATAGATATTTCTGGCTATTGTTGGTGTCCAAAGCTGAAACAACTCCAGTGGTGG	846
QY	1053	TAAATTTCTTAGAATTTCTATATGAAGAAATGATGAATCTGAGGTGGAAATTTATTCATGT	1112
Db	847	TAAATTTCTTAGAATTTCTATATGAAGAAATGATGAATCTGAGGTGGAAATTTATTCATGT	906
QY	1113	TACATCCCTATGTTGAAACAAAGGAGGAGATTCATTTCCGTTATCTTAAACAGGTAC	1172
Db	907	TACATCCCTATGTTGAAACAAAGGAGGAGATTCATTTCCGTTATCTTAAACAGGTAC	966
QY	1173	AGCAATCTCTAAAGTCACTTTTAAAGATGTCAGAAATATGATGATGCTGAAGGAAGAT	1232
Db	967	AGCAATCTCTAAAGTCACTTTTAAAGATGTCAGAAATATGATGATGCTGAAGGAAGAT	1026
QY	1233	CATAGATGTCATAGATGAAGAACTAATTCAACTTTGAGATTCATTTTGAAGAGTTGA	1292
Db	1027	CATAGATGTCATAGATGAAGAACTAATTCAACTTTGAGATTCATTTTGAAGAGTTGA	1086
QY	1293	ATATATTGCCAGAGCTGGATGGACTCTGAGGGAATAATGCTTGGTCCATCTACTAGA	1352
Db	1087	ATATATTGCCAGAGCTGGATGGACTCTGAGGGAATAATGCTTGGTCCATCTACTAGA	1146
QY	1353	TGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAATTTATTTATCCCACTAGA	1412
Db	1147	TGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAATTTATTTATCCCACTAGA	1206
QY	1413	AGATGATGTTATGGAAGGAGAGACTCAATGAGTCAGTGCCTGATTTCTGAGCGCCACT	1472
Db	1207	AGATGATGTTATGGAAGGAGAGACTCAATGAGTCAGTGCCTGATTTCTGAGCGCCACT	1266
QY	1473	ANTTATCTATGAAGAAACAAcAGACATCTGGATTAATATCCATGACATCTTTCTATGTTTT	1532
Db	1267	ANTTATCTATGAAGAAACAAcAGACATCTGGATTAATATCCATGACATCTTTCTATGTTTT	1326
QY	1533	TCCCCAAGTCAcGAGAGGAATTCAGTTTTATTTTGGCTCTGAATCCAAACAGGTTT	1592
Db	1327	TCCCCAAGTCAcGAGAGGAATTCAGTTTTATTTTGGCTCTGAATCCAAACAGGTTT	1386
QY	1593	CCGTCATTTATACAAATACATCTATTTTAAAGGAAGCAATATTAACCATCCAGTGG	1652
Db	1387	CCGTCATTTATACAAATACATCTATTTTAAAGGAAGCAATATTAACCATCCAGTGG	1446
QY	1653	TGGGCTGCCTCCAAAGTGATTTCAAGTGTCTATCAAAAGGAGGATAGCAATTAACAG	1712

Db	1447	TGGGCTGCCTCTCCAAAGTGATTTCAAGTGTCTCTATCAAAAGGAGGATAGCAATTAACCAG	1506
QY	1713	TGCTGAATGGGAAGTCTTTCGCGGCATGGATCTAATATATCCAAAGTTGATGAAGTCAGAG	1772
Db	1507	TGCTGAATGGGAAGTCTTTCGCGGCATGGATCTAATATATCCAAAGTTGATGAAGTCAGAG	1566
QY	1773	GCTGTATATATTTTGAAGGCACAAAGACTCCCTTTAGAGCATCAGCTCTAGTAGTCAG	1832
Db	1567	GCTGTATATATTTTGAAGGCACAAAGACTCCCTTTAGAGCATCAGCTCTAGTAGTCAG	1626
QY	1833	TTACGTAAATCCTCGAGAGGTGAcAAGGCTGACTGACCGTGCTACTCAcATTCTTGCTG	1892
Db	1627	TTACGTAAATCCTCGAGAGGTGAcAAGGCTGACTGACCGTGCTACTCAcATTCTTGCTG	1686
QY	1893	CATCAGTCAGACATCTGTGACTCTTTTATAAGTAAGTATAGTAACAGAGAATCCACATG	1952
Db	1687	CATCAGTCAGACATCTGTGACTCTTTTATAAGTAAGTATAGTAACAGAGAATCCACATG	1746
QY	1953	TGCTGCCCTTTTACAAGCTATCAAGTCTGGAAGATGACCCAACTTGCAAAACAAGGAATTT	2012
Db	1747	TGCTGCCCTTTTACAAGCTATCAAGTCTGGAAGATGACCCAACTTGCAAAACAAGGAATTT	1806
QY	2013	TTGGGCCACCAATTTTGGATTCAGCAGGTCTCTTCTGACTATCTCTCCAGAAATTTT	2072
Db	1807	TTGGGCCACCAATTTTGGATTCAGCAGGTCTCTTCTGACTATCTCTCCAGAAATTTT	1866
QY	2073	CTCTTTTGAAGTACTACTGGATTTTCAATGTATGGATGCTCTACAAGCCTCATGATCT	2132
Db	1867	CTCTTTTGAAGTACTACTGGATTTTCAATGTATGGATGCTCTACAAGCCTCATGATCT	1926
QY	2133	ACAGCTGGAAAGAAATATCTTACTGTGTGTTTCAATATATGGTGGTCTCAGGTGCGATT	2192
Db	1927	ACAGCTGGAAAGAAATATCTTACTGTGTGTTTCAATATATGGTGGTCTCAGGTGCGATT	1986
QY	2193	GGTGAATAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATATACCTAGACCTCTCAGG	2252
Db	1987	GGTGAATAATCGATTTTAAAGGAGTCAAGTATTTCCGCTTGAATATACCTAGACCTCTCAGG	2046
QY	2253	TTATGTGGTGTAGTAGACACACAGGGGATCTGTCTACCGAGGGCTTAAATTTGAAG	2312
Db	2047	TTATGTGGTGTAGTAGACACACAGGGGATCTGTCTACCGAGGGCTTAAATTTGAAG	2106
QY	2313	CGCTTTTAAATATATAAATGGTCAAAATAGAAATTTGACGATCAGGTGGAGGACTCCAATA	2372
Db	2107	CGCTTTTAAATATATAAATGGTCAAAATAGAAATTTGACGATCAGGTGGAGGACTCCAATA	2166
QY	2373	TCTAGCTTCTCGATATGATTTTCACTTGAATCTGTGGGCATCCACGGCTGGTCTTA	2432
Db	2167	TCTAGCTTCTCGATATGATTTTCACTTGAATCTGTGGGCATCCACGGCTGGTCTTA	2226
QY	2433	TGGAGGATACCTCTCCCTGATGGCATTAATGACAGAGTTCAGATATCTTTCAGGGTTGCTAT	2492
Db	2227	TGGAGGATACCTCTCCCTGATGGCATTAATGACAGAGTTCAGATATCTTTCAGGGTTGCTAT	2286
QY	2493	TGCTGGGGCCCCAGTCTACTCTGTGGATCTTCTATGATACAGGATACCGAAACGTTATAT	2552
Db	2287	TGCTGGGGCCCCAGTCTACTCTGTGGATCTTCTATGATACAGGATACCGAAACGTTATAT	2346
QY	2553	GGGTCAcCCTGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCCATCGCAGCAGA	2612
Db	2347	GGGTCAcCCTGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCCATCGCAGCAGA	2406
QY	2613	AAAGTTCCTCTGGAACCAAAATCGTTTACTGCTCTTACATGTTTCTTGGATGAGAATGT	2672
Db	2407	AAAGTTCCTCTGGAACCAAAATCGTTTACTGCTCTTACATGTTTCTTGGATGAGAATGT	2466
QY	2673	CCATTTTGCATACACAGTATATTACTGAGTTTTTTTAGTGGGCTGGAAAGCCATATGA	2732
Db	2467	CCATTTTGCATACACAGTATATTACTGAGTTTTTTTAGTGGGCTGGAAAGCCATATGA	2526
QY	2733	TTTACAGATCTATCTCCTCAGGAGAGACACAGCATTAAGAGTTCTCTGAATCGGAGAACATTA	2792

Db 2527 TTTACAGATCTATCCTCAGGAGACACAGCATAGAGTTCTGTAATCGGAGAACATTA 2586
Qy 2793 TGAACGCATCTTTTGCACCTACCTTCAAGAAACCTTGGATCAGCTATTTGCTCTAAA 2852
Db 2587 TGAATGCATCTTTTGCACCTACCTTCAAGAAACCTTGGATCAGCTATTTGCTCTAAA 2646
Qy 2853 AGTGATAT 2860
Db 2647 AGTGATAT 2654

RESULT 11
US-10-956-157-2177
; Sequence 2177, Application US/10956157
; Publication No. US20050118625A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William
; TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
; FILE OF INVENTION: HUMAN OSTEOARTHRITIS AND HUMAN PROTEASES
; FILE REFERENCE: 031896-043000 (AM 101081)
; CURRENT APPLICATION NUMBER: US/10/956,157
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 319805
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2177
; LENGTH: 2830
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-956-157-2177

Query Match 68.4%; Score 2135; DB 9; Length 2830;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2185; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 AAGTCTTAAGCCTCGAGCCCAAGCCGCTGCTACTGCGCCGCTGCTTACTAGTCGG 60
Db 1 AAGTCTTAAGCCTCGAGCCCAAGCCGCTGCTACTGCGCCGCTGCTTACTAGTCGG 60
Qy 61 CGTTCGCGCGCTGGGTTGTCAACCGCGCGCGCGCGAGAGGCCACTGCAACACAGACCG 120
Db 61 CGTTCGCGCGCTGGGTTGTCAACCGCGCGCGCGAGAGGCCACTGCAACACAGACCG 120
Qy 121 GAGTGGAGCGCGCGAGCATGAAGCGCGCGAGCCCGCTCCATAGCGCACTCGGAGCGG 180
Db 121 GAGTGGAGCGCGCGAGCATGAAGCGCGCGAGCCCGCTCCATAGCGCACTCGGAGCGG 180
Qy 181 TCCGCGCGCGCGCGGAGGAAATGCAACATGGCAGCAGCAATGGAAACAGAACAG 240
Db 181 TCCGCGCGCGCGCGGAGGAAATGCAACATGGCAGCAGCAATGGAAACAGAACAG 240
Qy 241 CTGGGTGTTGAGATATTTGAACTGCGGACTGTGAGGAGAAATTTGAATCAGAGATCGG 300
Db 241 CTGGGTGTTGAGATATTTGAACTGCGGACTGTGAGGAGAAATTTGAATCAGAGATCGG 300
Qy 301 CCTAAATTGGAGCCTTTTATGTTGAGCGGTATTCCTGAGTCAAGCTTAAAGAGCTGTT 360
Db 301 CCTAAATTGGAGCCTTTTATGTTGAGCGGTATTCCTGAGTCAAGCTTAAAGAGCTGTT 360
Qy 361 GCCGATACAGAAATATCATGCTACATGATGGCTTAAGGACACACATGATTTTCATGTTT 420
Db 361 GCCGATACAGAAATATCATGCTACATGATGGCTTAAGGACACACATGATTTTCATGTTT 420
Qy 421 GTGAAGAGGATGATCCAGATGACCTCATTCAGACAGAACTATACCTTGCATGCT 480
Db 421 GTGAAGAGGATGATCCAGATGACCTCATTCAGACAGAACTATACCTTGCATGCT 480
Qy 481 GGTGAGAACAGAAATATACCTGTTTATTCGAAATTCGAAACTATCATAGAGCA 540
Db 481 GGTGAGAACAGAAATATACCTGTTTATTCGAAATTCGAAACTATCATAGAGCA 540
Qy 541 GCAGTCTAATGCTCTTGGAGAGCCTCTTTTGGATCTTTTTCAGGCAACCTGGACTAT 600

Db 541 GCAGTCTAATGCTCTCTTTGGAAGCCTCTTTTGGATCTTTTTCAGGCAACACTGGACTAT 600
Qy 601 GGAATGTAATCTCGAGAGAGAGAACTATTTAAGAGAGAGAGAAACGATTTGGACAGTCGGA 660
Db 601 GGAATGTAATCTCGAGAGAGAGAACTATTTAAGAGAGAGAGAAACGATTTGGACAGTCGGA 660
Qy 661 ATTGCTCTTTACGATTTACCAAGAGAGTGGAACTATTTCTGTTTCAAGCCGCTAGTGA 720
Db 661 ATTGCTCTTTACGATTTACCAAGAGAGTGGAACTATTTCTGTTTCAAGCCGCTAGTGA 720
Qy 721 ATTTATCACGTAAGATGGAGGCCCAAGGATTTAGCAAAACCTTTTAAGGCCCAAT 780
Db 721 ATTTATCACGTAAGATGGAGGCCCAAGGATTTAGCAAAACCTTTTAAGGCCCAAT 780
Qy 781 CTAGTGGAAACTAGTTGTCTCCAAACATACGATGGATCCAAATTTATGCCCCTGATCCA 840
Db 781 CTAGTGGAAACTAGTTGTCTCCAAACATACGATGGATCCAAATTTATGCCCCTGATCCA 840
Qy 841 GACTGGATTTGCTTTTATACATAGCAACGATATTTGGATATCTTAACATCGTAACAGAGAA 900
Db 841 GACTGGATTTGCTTTTATACATAGCAACGATATTTGGATATCTTAACATCGTAACAGAGAA 900
Qy 901 GAAAGGAGACTCTACTTATGTGCAATGAGCTAGCAACATGGAAGAGATGCCAGATCA 960
Db 901 GAAAGGAGACTCTACTTATGTGCAATGAGCTAGCAACATGGAAGAGATGCCAGATCA 960
Qy 961 GCTGAGTCTGCTACCTTTGTTCTCCAAGAGAAATTTGATAGATATTTCTGGCTATTGGTGG 1020
Db 961 GCTGAGTCTGCTACCTTTGTTCTCCAAGAGAAATTTGATAGATATTTCTGGCTATTGGTGG 1020
Qy 1021 TGTCCAAAGCTGAAACAACTCCCGTGGTGTAAATTTCTAGAAATTTCTATATGAAGAA 1080
Db 1021 TGTCCAAAGCTGAAACAACTCCCGTGGTGTAAATTTCTAGAAATTTCTATATGAAGAA 1080
Qy 1081 AATCATGAATCTGAGGTGGAAATTTTATCTATCATCTGATCCCTATGTTGGAAACAGAGG 1140
Db 1081 AATCATGAATCTGAGGTGGAAATTTTATCTATCATCTGATCCCTATGTTGGAAACAGAGG 1140
Qy 1141 GCAGATTCATTCGTTATCTTAAACAGGTACAGCAATCTTAAAGTCACTTTTAAAGTG 1200
Db 1141 GCAGATTCATTCGTTATCTTAAACAGGTACAGCAATCTTAAAGTCACTTTTAAAGTG 1200
Qy 1201 TCAGAAATAATGATGATGCTGAAGAGGATCATAGATGTCATAGATAAGAACTAATTT 1260
Db 1201 TCAGAAATAATGATGATGCTGAAGAGGATCATAGATGTCATAGATAAGAACTAATTT 1260
Qy 1261 CAACCTTTTGAAGATTTCTATTTGAAGGATTTGAATATTTGCGCAGAGCTGGATGACTCCT 1320
Db 1261 CAACCTTTTGAAGATTTCTATTTGAAGGATTTGAATATTTGCGCAGAGCTGGATGACTCCT 1320
Qy 1321 GAGGGAATAATGCTTTGGTCCATCTTACTAGATCGCTCCGAGCTCGCCTACAGATAGTG 1380
Db 1321 GAGGGAATAATGCTTTGGTCCATCTTACTAGATCGCTCCGAGCTCGCCTACAGATAGTG 1380
Qy 1381 TTGATCTCACTGAAATTTATTTATCCAGTAGAAGATGATGTTTATGGAAGGAGAGACTC 1440
Db 1381 TTGATCTCACTGAAATTTATTTATCCAGTAGAAGATGATGTTTATGGAAGGAGAGACTC 1440
Qy 1441 ATTGAGTCAGTCTGATTTCTGAGCGCACTAAATTTCTATGAAGAAACAAACAGACATC 1500
Db 1441 ATTGAGTCAGTCTGATTTCTGAGCGCACTAAATTTCTATGAAGAAACAAACAGACATC 1500
Qy 1501 TGGATAATATCATGACATCTTTTATGTTTCCCAAGTCCAGAGAGAAATTTGAG 1560
Db 1501 TGGATAATATCATGACATCTTTTATGTTTCCCAAGTCCAGAGAGAAATTTGAG 1560
Qy 1561 TTTATTTTCCCTCTGATGCAAAACAGGTTTCCGTCATTTATACAAATTTACATCTATT 1620
Db 1561 TTTATTTTCCCTCTGATGCAAAACAGGTTTCCGTCATTTATACAAATTTACATCTATT 1620
Qy 1621 TTAAGGAAGAGAAATATAAGATCAGTGGTGGCTGCTGCTGAGTGATTTCAAG 1680
Db 1621 TTAAGGAAGAGAAATATAAGATCAGTGGTGGCTGCTGCTGAGTGATTTCAAG 1680


```

Db 1201 TCAGAAATATGATTGATGCTGAAGGAAGATCATAGATGTCTATAGATAAGAACTAAT 1260
QY CAACCTTTTGAAGATCTTATTTGAAGGAGTTGAATATATTTGCCAGAGCTGGAATGACCTCT 1320
Db 1261 CAACCTTTTGAAGATCTTATTTGAAGGAGTTGAATATATTTGCCAGAGCTGGAATGACCTCT 1320
QY GAGGAAATATGCTTGGTCCATCTTACTAGATCCGCTCCCAAGCTGCGCTACAGATAGTG 1380
Db 1321 GAGGAAATATGCTTGGTCCATCTTACTAGATCCGCTCCCAAGCTGCGCTACAGATAGTG 1380
QY TTGATCTCACTGAATATTTATTTCCAGTAGAAGATGATGATGATGATGGAAGCAGACTC 1440
Db 1381 TTGATCTCACTGAATATTTATTTCCAGTAGAAGATGATGATGATGATGGAAGCAGACTC 1440
QY ATTGAGTCAGTGCCTGATCTGTGAGCCCACTAATTTATCTATGAAAGAAACAACAGACATC 1500
Db 1441 ATTGAGTCAGTGCCTGATCTGTGAGCCCACTAATTTATCTATGAAAGAAACAACAGACATC 1500
QY TGGATAAATATCCATGACATCTTCTGATTTTCCCAAGTCCAGAGGAAATTTGAG 1560
Db 1501 TGGATAAATATCCATGACATCTTCTGATTTTCCCAAGTCCAGAGGAAATTTGAG 1560
QY TTTATTTTGGCTCTGAATGCAAAACAGGTTTCGGTCAATTTATACAAAATTTACATCTATT 1620
Db 1561 TTTATTTTGGCTCTGAATGCAAAACAGGTTTCGGTCAATTTATACAAAATTTACATCTATT 1620
QY TTAAGGAAAGCAAAATATATAACGATCCAGTGGTGGCTGCTGCTCCCAAGTATTTCAAG 1680
Db 1621 TTAAGGAAAGCAAAATATATAACGATCCAGTGGTGGCTGCTGCTCCCAAGTATTTCAAG 1680
QY TGTCCTATCAAGAGAGAGATAGCAATTTACAGTGGTGGCTGCTGCTCCCAAGTATTTCAAG 1740
Db 1681 TGTCCTATCAAGAGAGAGATAGCAATTTACAGTGGTGGCTGCTGCTCCCAAGTATTTCAAG 1740
QY GGAATCTAATATCAAGAGTGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
Db 1741 GGAATCTAATATCAAGAGTGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
QY TCCCTTTAGAGATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1860
Db 1801 TCCCTTTAGAGATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1860
QY CTGACTGACCGTGGCTACTCATTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1920
Db 1861 CTGACTGACCGTGGCTACTCATTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1920
QY AGTAAGTATAGTAACCAAGAAATCCACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1980
Db 1921 AGTAAGTATAGTAACCAAGAAATCCACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1980
QY GAAGATGACCAACTTGCMAAACAAGAAATTTTGGGCCACCAATTTTGGATTCAGCAGGT 2040
Db 1981 GAAGATGACCAACTTGCMAAACAAGAAATTTTGGGCCACCAATTTTGGATTCAGCAGGT 2040
QY CCTCTCTGACTATACCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGATTTACA 2100
Db 2041 CCTCTCTGACTATACCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGATTTACA 2100
QY TTGATGGATGCTCTCAAGCCTCATGATCTACAGCCTGGAAAGAAATATCTTACTGTG 2160
Db 2101 TTGATGGATGCTCTCAAGCCTCATGATCTACAGCCTGGAAAGAAATATCTTACTGTG 2160
QY CTGTTTCAATATGTTGGTGC 2179
Db 2161 CTGTTTCAATATGTTGGTGC 2179

```

RESULT 13
 US-10-982-512-8
 ; Sequence 8, Application US/10982512
 ; Publication No. US20050059081A1
 ; GENERAL INFORMATION:

; APPLICANT: Qi, Steve
 ; APPLICANT: Akinsanya, Karen
 ; APPLICANT: Riviere, Pierre
 ; APPLICANT: Junien, Jean-Louis
 ; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
 ; FILE REFERENCE: 70669
 ; CURRENT APPLICATION NUMBER: US/10/982,512
 ; PRIOR FILING DATE: 2004-11-05
 ; PRIOR APPLICATION NUMBER: US/09/976,674
 ; PRIOR FILING DATE: 2001-10-12
 ; PRIOR APPLICATION NUMBER: US 60/240,117
 ; NUMBER OF SEQ ID NOS: 61
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 8
 ; LENGTH: 4523
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-982-512-8

Query Match 68.2%; Score 2128; DB 9; Length 4523;
 Best Local Similarity 99.9%; Pred No. 0;
 Matches 2178; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 AAGTGTCTAAGCCTCGAGGCCAAGCCGCTGCTACTCTGCCGCCGCTCTCTTAGTGCGG 60
Db 1 AAGTGTCTAAGCCTCGAGGCCAAGCCGCTGCTACTGCCGCCGCTCTCTTAGTGCGG 60
QY 61 CGTTGCCGCCCTGGTGTGTCAACCGCGCGCGCGAGGAGCACTGCAACACGAGACCG 120
Db 61 CGTTGCCGCCCTGGTGTGTCAACCGCGCGCGCGAGGAGCACTGCAACACGAGACCG 120
QY 121 GAGTGAGCGCGCGCAGCATGAAGCGCGCGCGCGCTCCATAGCGCAGCTCGGACCG 180
Db 121 GAGTGAGCGCGCGCAGCATGAAGCGCGCGCGCGCTCCATAGCGCAGCTCGGACCG 180
QY 181 TCGGCGCGCGCGCGGGAAGAGAAATGCAACATGCGCAGCAATGGAACAGAAACAG 240
Db 181 TCGGCGCGCGCGCGGGAAGAGAAATGCAACATGCGCAGCAATGGAACAGAAACAG 240
QY 241 CTGGGTGTGAGATATTTGAAACTGCGGACTGTGAGGAGAAATTTGAAATCAGAGGATCG 300
Db 241 CTGGGTGTGAGATATTTGAAACTGCGGACTGTGAGGAGAAATTTGAAATCAGAGGATCG 300
QY 301 CCTAAATTTGAGCGCTTTTATTTAGTCGGGTATTTCTGGAGTCAGCTTAAAGCTGCTT 360
Db 301 CCTAAATTTGAGCGCTTTTATTTAGTCGGGTATTTCTGGAGTCAGCTTAAAGCTGCTT 360
QY 361 GCGGATACAGAAATATCATGGCTACATGATGACCTCATTCAGACAGAAATCTATACCTGCCATGCT 420
Db 361 GCGGATACAGAAATATCATGGCTACATGATGACCTCATTCAGACAGAAATCTATACCTGCCATGCT 420
QY 421 GTGAAGAGGAATGATCCAGATGGACCTCATTCAGACAGAAATCTATACCTGCCATGCT 480
Db 421 GTGAAGAGGAATGATCCAGATGGACCTCATTCAGACAGAAATCTATACCTGCCATGCT 480
QY 481 GGTGAGACAGAGAAATACACTGTTTATTTCTGAAATTTCCCAAACTATCAATAGACA 540
Db 481 GGTGAGACAGAGAAATACACTGTTTATTTCTGAAATTTCCCAAACTATCAATAGACA 540
QY 541 GCATCTTAATGCTCTCTTGGAAAGCCTCTTTGGATCTTTTTCAGGCAACACTGACTAT 600
Db 541 GCATCTTAATGCTCTCTTGGAAAGCCTCTTTGGATCTTTTTCAGGCAACACTGACTAT 600
QY 601 GGAATGATTTCTCGAAGAAAGAACTATTTAAGAGAAAGAAACGCAATTTGGAAACAGTCGGA 660
Db 601 GGAATGATTTCTCGAAGAAAGAACTATTTAAGAGAAAGAAACGCAATTTGGAAACAGTCGGA 660
QY 661 ATTGCTCTTACGATTATCAACAGGAAGTGAACATTTCTGTTTCAAGCCGCTAGTGGA 720
Db 661 ATTGCTCTTACGATTATCAACAGGAAGTGAACATTTCTGTTTCAAGCCGCTAGTGGA 720
QY 721 ATTTATCACTAAAGATCGAGGCGCACAGGATTTACGCAACACCTTTTAAGCCCAAT 780

```


Db	721	ATTTATACGTAATAAGATGGAGGCGCAAGAGATTTACGCAACACCTTTAAGGCCCAAT	780	Db	1801	TCCTCTTTAGAGCATCACCTGTACGTAGTCAGTTACGTAAATCCTCGGAGAGTGACAAGG	1860
Qy	781	CTAGTGGAACTAGTGTGTCCTCAACATACGGATGGATCCAAAATTAATGCCCCCTGATCCA	840	Qy	1861	CTGACTGACCGTGGCTACTCTACATTTCTGTCATCAGTCAGCAGCTGTCGACTTCTTTTATA	1920
Db	781	CTAGTGGAACTAGTGTGTCCTCAACATACGGATGGATCCAAAATTAATGCCCCCTGATCCA	840	Db	1861	CTGACTGACCGTGGCTACTCTACATTTCTGTCATCAGTCAGTCAGCAGCTGTCGACTTCTTTATA	1920
Qy	841	GACTGGATTGCTTTTATACATAGCAAGCATATTGGGATATCTAAACATCGTAACCAAGAGAA	900	Qy	1921	AGTAAGTATAGTAAACCAAGAGAAATCCACACTGTGTGTGTCCTTTTCAAGTATCAAGTCCT	1980
Db	841	GACTGGATTGCTTTTATACATAGCAAGCATATTGGGATATCTAAACATCGTAACCAAGAGAA	900	Db	1921	AGTAAGTATAGTAAACCAAGAGAAATCCACACTGTGTGTGTCCTTTTCAAGTATCAAGTCCT	1980
Qy	901	GAAAGAGACTCACTTATGTGCAATAGCTAGCTAGCAATGGAAGAGATGCCAGATCA	960	Qy	1981	GAAGATGACCCCACTTGGCAAAACAAAGAAATTTTGGGCCACCAATTTTGGATTCAGCAGGT	2040
Db	901	GAAAGAGACTCACTTATGTGCAATAGCTAGCTAGCAATGGAAGAGATGCCAGATCA	960	Db	1981	GAAGATGACCCCACTTGGCAAAACAAAGAAATTTTGGGCCACCAATTTTGGATTCAGCAGGT	2040
Qy	961	GCTGGAGTGCCTACCTTTGTTCTCCAGAGAAATTTGATAGATATTTCTGGCTATTGGTGG	1020	Qy	2041	CCTCTTCTGACTATACCTCCAGAAATTTTCTTTTGAAGTACTACTCTGGATTTACA	2100
Db	961	GCTGGAGTGCCTACCTTTGTTCTCCAGAGAAATTTGATAGATATTTCTGGCTATTGGTGG	1020	Db	2041	CCTCTTCTGACTATACCTCCAGAAATTTTCTTTTGAAGTACTACTCTGGATTTACA	2100
Qy	1021	TGTCCAAAAGCTGAAACAACTCCAGTGTGTGTAATAATCTTAGAATTTCTATAGAGAA	1080	Qy	2101	TTGTATGGGATGCTCTACAAGCTCATGATCTACAGCTCGAAGAAATATCCTACTGTG	2160
Db	1021	TGTCCAAAAGCTGAAACAACTCCAGTGTGTGTAATAATCTTAGAATTTCTATAGAGAA	1080	Db	2101	TTGTATGGGATGCTCTACAAGCTCATGATCTACAGCTCGAAGAAATATCCTACTGTG	2160
Qy	1081	AATGATGAATCTGAGTGGAAATTAATCATGTTACATCCCTATGTTGGAAACAAGGAG	1140	Qy	2161	CTGTTTCATATATGTTGGTC	2179
Db	1081	AATGATGAATCTGAGTGGAAATTAATCATGTTACATCCCTATGTTGGAAACAAGGAG	1140	Db	2161	CTGTTTCATATATGTTGGTC	2179
Qy	1141	GCAGATTCAATTCGGTTATCCTTAAACACAGGTACAGCAATCTTAAAGTCATTTTAAGATG	1200	RESULT 14			
Db	1141	GCAGATTCAATTCGGTTATCCTTAAACACAGGTACAGCAATCTTAAAGTCATTTTAAGATG	1200	US-09-976-674-20			
Qy	1201	TGAGAAATATGATGCTCAAGAGAGGATCATAGATGTCATAGTAAGAACTAAAT	1260	; Sequence 20, Application US/09976674			
Db	1201	TGAGAAATATGATGCTCAAGAGAGGATCATAGATGTCATAGTAAGAACTAAAT	1260	; Patent No. US20020115843A1			
Qy	1261	CRACCTTTTGAGATTTCTATTGAAGAGTTGAATATATTGCCAGAGCTGGATGACTCT	1320	; GENERAL INFORMATION:			
Db	1261	CRACCTTTTGAGATTTCTATTGAAGAGTTGAATATATTGCCAGAGCTGGATGACTCT	1320	; APPLICANT: Qi, Steve			
Qy	1321	GAGGAAATATGCTTGGTCCATCTCTAGATGCTCCAGACTCCAGACTGCTGATGATG	1380	; APPLICANT: Akinsanya, Karen			
Db	1321	GAGGAAATATGCTTGGTCCATCTCTAGATGCTCCAGACTCCAGACTGCTGATGATG	1380	; APPLICANT: Riviere, Pierre			
Qy	1381	TTGATCTCACTGAAATTTATCCAGTAGAAGATGATGTTATGGAAGCGAGAGACT	1440	; APPLICANT: Junish, Jean-Louis			
Db	1381	TTGATCTCACTGAAATTTATCCAGTAGAAGATGATGTTATGGAAGCGAGAGACT	1440	; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV			
Qy	1441	ATTGAGTCAGTGCCTGATCTGTGAGCCCACTAATTTATCTATGAAGAAACAAACAGATC	1500	; FILE REFERENCE: 70669			
Db	1441	ATTGAGTCAGTGCCTGATCTGTGAGCCCACTAATTTATCTATGAAGAAACAAACAGATC	1500	; CURRENT APPLICATION NUMBER: US/09/976,674			
Qy	1501	TGGATAAATATCCATGACATCTTTCAATGTTTTTCCCAAAGTCACGAAGAGAAATGAG	1560	; CURRENT FILING DATE: 2001-10-12			
Db	1501	TGGATAAATATCCATGACATCTTTCAATGTTTTTCCCAAAGTCACGAAGAGAAATGAG	1560	; PRIOR APPLICATION NUMBER: US 60/240,117			
Qy	1561	TTTATTTTGGCTCTGAAATCAAAACAGGTTTCGGTCAATTTATACAAATTTACATCTATT	1620	; PRIOR FILING DATE: 2000-10-12			
Db	1561	TTTATTTTGGCTCTGAAATCAAAACAGGTTTCGGTCAATTTATACAAATTTACATCTATT	1620	; NUMBER OF SEQ ID NOS: 61			
Qy	1621	TTAAGGAAAGCAAAATATAACGATCCAGTGTGGGCTGCTGCTCCAAAGTATTTCAAG	1680	; SOFTWARE: PatentIn version 3.1			
Db	1621	TTAAGGAAAGCAAAATATAACGATCCAGTGTGGGCTGCTGCTCCAAAGTATTTCAAG	1680	; SEQ ID NO 20			
Qy	1681	TGTCTATCAAGAGGAGATAGCAATTAACAGTGTGTAATGGGAAGTTCTTGGCCGGCAT	1740	; LENGTH: 4676			
Db	1681	TGTCTATCAAGAGGAGATAGCAATTAACAGTGTGTAATGGGAAGTTCTTGGCCGGCAT	1740	; TYPE: DNA			
Qy	1741	GGATCTAATATCAAGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGGCAACCAAGAC	1800	; ORGANISM: Homo sapiens			
Db	1741	GGATCTAATATCAAGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGGCAACCAAGAC	1800	US-09-976-674-20			
Qy	1801	TCCTCTTTAGAGCATCACCTGTACGTAGTCAGTTACGTAAATCCTCGGAGAGTGACAAGG	1860	Query Match 68.2%; Score 2128; DB 3; Length 4676;			
Db				Best Local Similarity 99.9%; Pred. No. 0;			
				Matches 2178; Conservative 0; Mismatches 1; Indels 0; Gaps 0;			
Qy	1	AAGTGTCTAAAGCCTCCGAGGCCAAGCCGCTCTACTGCGCGCGCTGCTTCTTAGTGCGG	60	Qy	1	AAGTGTCTAAAGCCTCCGAGGCCAAGCCGCTCTACTGCGCGCGCTGCTTCTTAGTGCGG	60
Db	1	AAGTGTCTAAAGCCTCCGAGGCCAAGCCGCTCTACTGCGCGCGCTGCTTCTTAGTGCGG	60	Db	1	AAGTGTCTAAAGCCTCCGAGGCCAAGCCGCTCTACTGCGCGCGCTGCTTCTTAGTGCGG	60
Qy	61	CGTTTCGCCGCTGGTGTTCACCGCGCGCGCGCGAGGAGCCACTGCAACCCAGGACCG	120	Qy	61	CGTTTCGCCGCTGGTGTTCACCGCGCGCGCGCGAGGAGCCACTGCAACCCAGGACCG	120
Db	61	CGTTTCGCCGCTGGTGTTCACCGCGCGCGCGCGAGGAGCCACTGCAACCCAGGACCG	120	Db	61	CGTTTCGCCGCTGGTGTTCACCGCGCGCGCGCGAGGAGCCACTGCAACCCAGGACCG	120
Qy	121	GAGTGGAGGCGCGCGCAGCATGAAGCGCGCGCGCGCTCCATAGCGCAGCTCGGAGCGG	180	Qy	121	GAGTGGAGGCGCGCGCAGCATGAAGCGCGCGCGCGCTCCATAGCGCAGCTCGGAGCGG	180
Db	121	GAGTGGAGGCGCGCGCAGCATGAAGCGCGCGCGCGCTCCATAGCGCAGCTCGGAGCGG	180	Db	121	GAGTGGAGGCGCGCGCAGCATGAAGCGCGCGCGCGCTCCATAGCGCAGCTCGGAGCGG	180
Qy	181	TCCTCTTTAGAGTATTTGAAATCTGCGGACTGTGAGGAGAAATTTGAATTCACAGATCGG	240	Qy	181	TCCTCTTTAGAGTATTTGAAATCTGCGGACTGTGAGGAGAAATTTGAATTCACAGATCGG	240
Db	181	TCCTCTTTAGAGTATTTGAAATCTGCGGACTGTGAGGAGAAATTTGAATTCACAGATCGG	240	Db	181	TCCTCTTTAGAGTATTTGAAATCTGCGGACTGTGAGGAGAAATTTGAATTCACAGATCGG	240
Qy	241	CTGGGTGTTGAGATATTTGAAATCTGCGGACTGTGAGGAGAAATTTGAATTCACAGATCGG	300	Qy	241	CTGGGTGTTGAGATATTTGAAATCTGCGGACTGTGAGGAGAAATTTGAATTCACAGATCGG	300
Db	241	CTGGGTGTTGAGATATTTGAAATCTGCGGACTGTGAGGAGAAATTTGAATTCACAGATCGG	300	Db	241	CTGGGTGTTGAGATATTTGAAATCTGCGGACTGTGAGGAGAAATTTGAATTCACAGATCGG	300

;; PRIOR FILING DATE: 2000-10-12
;; NUMBER OF SEQ ID NOS: 61
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 20
;; LENGTH: 4676
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-10-982-512-20

Query Match 68.2%; Score 2128; DB 9; Length 4676;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 2178; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	1	AAGTCTAAAGCCTCCGAGGCCAAGGCGCTGCTACTGCGCGCTCTTCTAGTGGCG	60
DB	1	AAGTCTAAAGCCTCCGAGGCCAAGGCGCTGCTACTGCGCGCTCTTCTAGTGGCG	60
QY	61	CGTTCCGCGCTGGGTTGTACACGCGCGCCGCCGAGGAAGCCACTGCAACAGGACCG	120
DB	61	CGTTCCGCGCTGGGTTGTACACGCGCGCCGCCGAGGAAGCCACTGCAACAGGACCG	120
QY	121	GAGTGGAGCGCGCAGCATGAAGCGCGCGCAGGCCCGCTCCATAGCGCACGTGCGGACGG	180
DB	121	GAGTGGAGCGCGCAGCATGAAGCGCGCGCAGGCCCGCTCCATAGCGCACGTGCGGACGG	180
QY	181	TCGCGCGCGCGCGGGGAGGAAATGCAACATGGCAGCAGCAATGGAACAGAACAG	240
DB	181	TCGCGCGCGCGCGGGGAGGAAATGCAACATGGCAGCAGCAATGGAACAGAACAG	240
QY	241	CTGGGTGTGAGATATTTGAAACTGCGGACTGTGAGGAGAAATTTGAATCACAGGATCGG	300
DB	241	CTGGGTGTGAGATATTTGAAACTGCGGACTGTGAGGAGAAATTTGAATCACAGGATCGG	300
QY	301	CTTAATTTGGAGCCTTTTATGTTGAGCGGTATTCCTGGGTGACGTTAAAGAGTCGTT	360
DB	301	CTTAATTTGGAGCCTTTTATGTTGAGCGGTATTCCTGGGTGACGTTAAAGAGTCGTT	360
QY	361	GCGATACAGAAATATCATGGCTACATGATGGCTAAGGACCAACATGATTTCAATGTTT	420
DB	361	GCGATACAGAAATATCATGGCTACATGATGGCTAAGGACCAACATGATTTCAATGTTT	420
QY	421	GTGAAGAGGAATGATCCAGATGGACCTCATTACAGACAGAAATCTATTACCTTGCCATGTCT	480
DB	421	GTGAAGAGGAATGATCCAGATGGACCTCATTACAGACAGAAATCTATTACCTTGCCATGTCT	480
QY	481	GGTGAGAACAGAGAAATACACTGTTTTTATCTGAAATTTCCCAAACTATCAATAGACA	540
DB	481	GGTGAGAACAGAGAAATACACTGTTTTTATCTGAAATTTCCCAAACTATCAATAGACA	540
QY	541	GGAGTCTTAATGCTCTCTTGAAGCCTCTTTTGGATCTTTTTCAGGCAACACTGGACTAT	600
DB	541	GGAGTCTTAATGCTCTCTTGAAGCCTCTTTTGGATCTTTTTCAGGCAACACTGGACTAT	600
QY	601	GGATGTATTTCCAGAGAGAACTATTAAAGAGAAAGAAACGATTTGGACACTCGGA	660
DB	601	GGATGTATTTCCAGAGAGAACTATTAAAGAGAAAGAAACGATTTGGACACTCGGA	660
QY	661	ATTGCTCTTACGATTTACCAAGGAAGTGAACATTTCTGTTTCAAGCCGGTAGTGA	720
DB	661	ATTGCTCTTACGATTTACCAAGGAAGTGAACATTTCTGTTTCAAGCCGGTAGTGA	720
QY	721	ATTATACGTTAAAGATGGAGGCCACAGGATTTTACGCAACAACTTTTAAGGCCCAAT	780
DB	721	ATTATACGTTAAAGATGGAGGCCACAGGATTTTACGCAACAACTTTTAAGGCCCAAT	780
QY	781	CTAGTGGAAACTAGTGTGTCCTCAACATACGATGATCCAAATTTATGCCCTGCTGATCCA	840
DB	781	CTAGTGGAAACTAGTGTGTCCTCAACATACGATGATCCAAATTTATGCCCTGCTGATCCA	840
QY	841	GACTGGATGCTTTTATACATAGCAACGATATTTGGATATCTAAACATCGTAAACAGAGAA	900
DB	841	GACTGGATGCTTTTATACATAGCAACGATATTTGGATATCTAAACATCGTAAACAGAGAA	900

QY	901	GAAAGGAGACTCACTTATGTGCACAATGAGCTAGCCAAATGGAAGAAGATCCAGATCA	960
DB	901	GAAAGGAGACTCACTTATGTGCACAATGAGCTAGCCAAATGGAAGAAGATCCAGATCA	960
QY	961	GCTGGAGTCGCTACCTTTGTTCTCCAAAGAAATTTGATAGATATTTCTGGCTATTGGTGG	1020
DB	961	GCTGGAGTCGCTACCTTTGTTCTCCAAAGAAATTTGATAGATATTTCTGGCTATTGGTGG	1020
QY	1021	TGTCCTCAAAAGCTGAACCAACTCCAGTGGTGTAAATTTCTTAGAATTTCTATATGAAGA	1080
DB	1021	TGTCCTCAAAAGCTGAACCAACTCCAGTGGTGTAAATTTCTTAGAATTTCTATATGAAGA	1080
QY	1081	AATGATGAATCTGAGGTGGAAATTTATTCATGTTTACATCCCTATGTTGGAACAAGGAGG	1140
DB	1081	AATGATGAATCTGAGGTGGAAATTTATTCATGTTTACATCCCTATGTTGGAACAAGGAGG	1140
QY	1141	GCAGATTCATTTCCGTTATCTCTAAACAGGTACAGCAAAATCTTAAAGTCACTTTTAAAGATG	1200
DB	1141	GCAGATTCATTTCCGTTATCTCTAAACAGGTACAGCAAAATCTTAAAGTCACTTTTAAAGATG	1200
QY	1201	TCAGAAATTAATGATGCTGGAAGGAGATCATAGATGTCTATAGATAAGGAACTAATTT	1260
DB	1201	TCAGAAATTAATGATGCTGGAAGGAGATCATAGATGTCTATAGATAAGGAACTAATTT	1260
QY	1261	CAACCTTTTGAGATTTCTATTGGAAGGATTTGAATATATTGCCCAGAGCTGGATGGACTCT	1320
DB	1261	CAACCTTTTGAGATTTCTATTGGAAGGATTTGAATATATTGCCCAGAGCTGGATGGACTCT	1320
QY	1321	GAGGAAATATGCTTCGTCCTACTACTAGATCGCTCCAGACTCGCTACAGATAGTG	1380
DB	1321	GAGGAAATATGCTTCGTCCTACTACTAGATCGCTCCAGACTCGCTACAGATAGTG	1380
QY	1381	TTGATCTCACCTGAATTTATTTATCCAGTAGAAGATGATGTTATGGAAGGAGAGACTC	1440
DB	1381	TTGATCTCACCTGAATTTATTTATCCAGTAGAAGATGATGTTATGGAAGGAGAGACTC	1440
QY	1441	ATTGAGTCAGTGCCTGATTTCTGTAGCCCACTAATTTATCTATGAAGAAACACAGACATC	1500
DB	1441	ATTGAGTCAGTGCCTGATTTCTGTAGCCCACTAATTTATCTATGAAGAAACACAGACATC	1500
QY	1501	TGGATAAATPCCATGACATCTTTTCATGTTTTTCCCAAGTCAAGAGAGAAATGAG	1560
DB	1501	TGGATAAATPCCATGACATCTTTTCATGTTTTTCCCAAGTCAAGAGAGAAATGAG	1560
QY	1561	TTTATTTTGGCTCTGAATGCAAAACAGGTTTCCGTCATTTATACAAAATTTACATCTAT	1620
DB	1561	TTTATTTTGGCTCTGAATGCAAAACAGGTTTCCGTCATTTATACAAAATTTACATCTAT	1620
QY	1621	TTAAAGGAAAGCAATATAACGATCCAGTGGTGGGCTGCTCTCCAGATGATTTCAAG	1680
DB	1621	TTAAAGGAAAGCAATATAACGATCCAGTGGTGGGCTGCTCTCCAGATGATTTCAAG	1680
QY	1681	TGTCCTATCAAGAGGAGATAGCAATTTACAGTGGTGGTGAATGGAAAGTTCTTCGCGCGCAT	1740
DB	1681	TGTCCTATCAAGAGGAGATAGCAATTTACAGTGGTGGTGAATGGAAAGTTCTTCGCGCGCAT	1740
QY	1741	GGATCTAATATCCAAAGTTGATGAAGTCAAGAGGCTGGTATATTTTGAAGGCCACCAAGAC	1800
DB	1741	GGATCTAATATCCAAAGTTGATGAAGTCAAGAGGCTGGTATATTTTGAAGGCCACCAAGAC	1800
QY	1801	TCCCTTTTGAAGCATACCTGTACGTAGTACGTAAATCTCTGGAGAGGTGACAAGG	1860
DB	1801	TCCCTTTTGAAGCATACCTGTACGTAGTACGTAAATCTCTGGAGAGGTGACAAGG	1860
QY	1861	CTGACTGACCGTGGCTACTCACAATTTCTGTCATCAGTCAGCAGCTGTCTTTTATA	1920
DB	1861	CTGACTGACCGTGGCTACTCACAATTTCTGTCATCAGTCAGCAGCTGTCTTTTATA	1920
QY	1921	AGTAAGATAGTACCAAGAGAAATCCACACTGTGTGCTCCCTTTTCAAGCTATCAAGTCT	1980
DB	1921	AGTAAGATAGTACCAAGAGAAATCCACACTGTGTGCTCCCTTTTCAAGCTATCAAGTCT	1980
QY	1981	GAAGATGACCCCACTTCCAAAAACAAAGGAATTTTGGGCCACCACTTTTGGATTTCAGCAGGT	2040

Accession	Sequence	Length
1981	GAAGATGACCAACTTCGAAACAAAGGAATTTTGGGCCACCATTTTGGATTTACAGAGT	2040
2041	CCTCTTCCTGACTATCTCCTCCAGAAAATTTCTTTTGGAAAGTACTACTGGATTTTACA	2100
2041	CCTCTTCCTGACTATCTCCTCCAGAAAATTTCTTTTGGAAAGTACTACTGGATTTTACA	2100
2101	TTGTATGGATGCTCTACAGCCTCATGATCTACAGCCTGGGAAGAATATCCTACTGTG	2160
2101	TTGTATGGATGCTCTACAGCCTCATGATCTACAGCCTGGGAAGAATATCCTACTGTG	2160
2161	CTGTTTCATATATGGTGGTC	2179
2161	CTGTTTCATATATGGTGGTC	2179

Search completed: May 4, 2006, 10:45:48
Job time : 2199 secs

GenCore version 5.1.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 4, 2006, 09:26:31 ; Search time 1177 Seconds
(without alignments)
10795.511 Million cell updates/sec

Title: US-10-825-632-2

Perfect score: 3120

Sequence: 1 aagtgttaagcttcgagg.....agaattactataaaaaaaa 3120

Scoring table: **OLIGO_NUC**

Gapop 60.0 , Gapext 60.0

Searched: 9306428 seqs, 2036268586 residues

Word size: 8

Total number of hits satisfying chosen parameters: 18612072

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published Applications NA_New.*

1: /SIDSS/ptodata/2/pubpna/US08_NEW_PUB.seq1.*
2: /SIDSS/ptodata/2/pubpna/US06_NEW_PUB.seq.*
3: /SIDSS/ptodata/2/pubpna/US07_NEW_PUB.seq.*
4: /SIDSS/ptodata/2/pubpna/US08_NEW_PUB.seq.*
5: /SIDSS/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
6: /SIDSS/ptodata/2/pubpna/US09_NEW_PUB.seq.*
7: /SIDSS/ptodata/2/pubpna/US09_NEW_PUB.seq1.*
8: /SIDSS/ptodata/2/pubpna/US09_NEW_PUB.seq2.*
9: /SIDSS/ptodata/2/pubpna/US10_NEW_PUB.seq.*
10: /SIDSS/ptodata/2/pubpna/US10_NEW_PUB.seq1.*
11: /SIDSS/ptodata/2/pubpna/US10_NEW_PUB.seq2.*
12: /SIDSS/ptodata/2/pubpna/US10_NEW_PUB.seq3.*
13: /SIDSS/ptodata/2/pubpna/US10_NEW_PUB.seq4.*
14: /SIDSS/ptodata/2/pubpna/US11_NEW_PUB.seq.*
15: /SIDSS/ptodata/2/pubpna/US11_NEW_PUB.seq2.*
16: /SIDSS/ptodata/2/pubpna/US11_NEW_PUB.seq3.*
17: /SIDSS/ptodata/2/pubpna/US11_NEW_PUB.seq4.*
18: /SIDSS/ptodata/2/pubpna/US11_NEW_PUB.seq5.*
19: /SIDSS/ptodata/2/pubpna/US60_NEW_PUB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Length	ID	Description
1	2957	94.8	3143	17 US-11-151-601-19 Sequence 19, Appl
2	2547	81.6	2649	17 US-11-151-601-21 Sequence 21, Appl
3	118	3.8	1346	7 US-09-925-065A-669313 Sequence 669313,
4	112	3.6	609	7 US-09-925-065A-743558 Sequence 743558,
5	112	3.6	609	7 US-09-925-065A-743559 Sequence 743559,
6	73	2.3	612	7 US-09-925-065A-818446 Sequence 818446,
7	28	0.9	497	7 US-09-925-065A-482467 Sequence 482467,
8	26	0.8	26	10 US-10-310-914A-543495 Sequence 543495,
9	25	0.8	25	17 US-11-121-849-207313 Sequence 207313,
10	25	0.8	25	17 US-11-121-849-207314 Sequence 207314,
11	25	0.8	25	17 US-11-121-849-207315 Sequence 207315,
12	25	0.8	25	17 US-11-121-849-207316 Sequence 207316,
13	25	0.8	25	17 US-11-121-849-207317 Sequence 207317,
14	25	0.8	25	17 US-11-121-849-207318 Sequence 207318,

15	25	0.8	25	17	US-11-121-849-207319	Sequence 207319,
16	25	0.8	25	17	US-11-121-849-207320	Sequence 207320,
17	25	0.8	25	17	US-11-121-849-207321	Sequence 207321,
18	25	0.8	25	17	US-11-121-849-207322	Sequence 207322,
19	25	0.8	25	17	US-11-121-849-207323	Sequence 207323,
20	25	0.8	28	10	US-10-310-914A-543509	Sequence 543509,
21	23	0.7	23	10	US-10-310-914A-543503	Sequence 543503,
22	23	0.7	23	10	US-10-310-914A-543506	Sequence 543506,
23	23	0.7	1275	17	US-11-128-061-624	Sequence 624, App
24	23	0.7	1275	17	US-11-128-049-624	Sequence 624, App
25	22	0.7	22	10	US-10-310-914A-543494	Sequence 543494,
26	22	0.7	22	10	US-10-310-914A-543501	Sequence 543501,
27	22	0.7	22	10	US-10-310-914A-543507	Sequence 543507,
28	22	0.7	22	10	US-10-310-914A-543510	Sequence 543510,
29	21	0.7	21	10	US-10-310-914A-543510	Sequence 543510,
30	21	0.7	21	10	US-10-310-914A-543502	Sequence 543502,
31	21	0.7	21	10	US-10-310-914A-543508	Sequence 543508,
32	21	0.7	535	18	US-11-226-869-428	Sequence 543513,
33	21	0.7	611	7	US-09-925-065A-716183	Sequence 428, App
34	21	0.7	611	7	US-09-925-065A-716184	Sequence 716183,
35	21	0.7	976	18	US-11-096-568A-307	Sequence 716184,
36	21	0.7	1981	7	US-09-925-065A-699851	Sequence 307, App
37	21	0.7	1981	7	US-09-925-065A-699852	Sequence 699851,
38	21	0.7	2159	10	US-10-750-185-47863	Sequence 699852,
39	21	0.7	2159	10	US-10-750-623-47863	Sequence 47863, A
40	21	0.7	2788	9	US-10-505-928-476	Sequence 47863, A
41	21	0.7	2814	17	US-11-186-284-54	Sequence 476, App
42	21	0.7	2814	18	US-11-245-147-168	Sequence 54, Appl
43	21	0.7	171732	17	US-11-121-086-98	Sequence 168, Appl
44	21	0.7	191150	13	US-10-517-905-19	Sequence 98, Appl
45	20	0.6	20	10	US-10-310-914A-543499	Sequence 19, Appl
						Sequence 543499,

ALIGNMENTS

RESULT 1
US-11-151-601-19
; Sequence 19, Application US/11151601
; Publication No. US20060003413A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Olandt, Peter J.
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Curtis, Rory A. J.
; APPLICANT: Williamson, Mark
; APPLICANT: Welch, Nadine
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASE, PHOSPHATASE.
; FILE REFERENCE: MPI00-054PINCPIOMNIDVIM,
; CURRENT APPLICATION NUMBER: US/11/151,601
; CURRENT FILING DATE: 2005-06-13
; PRIOR APPLICATION NUMBER: US 10/170,789
; PRIOR FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: US 09/797,039
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: PCT/US01/06525
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/186,061
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 09/882,166
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: PCT/US01/19269
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/212,078
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 09/934,406
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: PCT/US01/26052
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/226,740
; PRIOR FILING DATE: 2000-08-21

```
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: BlastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 3143
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (229) ... (2874)
US-11-151-601-19

Query Match          94.8%; Score 2957; DB 17; Length 3143;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3107; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 11 GCCTCCGAGGCCAAGGCCGCTGCTACTGCGCGCGCTGCTTCTTAGTGCCTGCTTTCGCGC 70
DB 26 GCCTCCGAGGCCAAGGCCGCTGCTACTGCGCGCGCTGCTTCTTAGTGCCTGCTTTCGCGC 85

QY 71 CTGGGTTGTCAACCGGCGCGCGCGCGAGGAGCCACTGCACACGAGCCGCGAGTGGAGGC 130
DB 86 CTGGGTTGTCAACCGGCGCGCGCGCGAGGAGCCACTGCACACGAGCCGCGAGTGGAGGC 145

QY 131 GCGCCAGCATGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 190
DB 146 GCGCCAGCATGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 205

QY 191 GCCCGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 250
DB 206 GCCCGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 265

QY 251 AGATATTGAACTCGGAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 310
DB 266 AGATATTGAACTCGGAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 325

QY 311 AGCCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAGGCTGCTTCCGCGATACCA 370
DB 326 AGCCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAGGCTGCTTCCGCGATACCA 385

QY 371 GAAATATCATGCTACATGAGGCTTAAGGCAACCATGATTTTACCTTTGCCATGCTCTGGTGAACA 430
DB 386 GAAATATCATGCTACATGAGGCTTAAGGCAACCATGATTTTACCTTTGCCATGCTCTGGTGAACA 445

QY 431 ATGATCCAGATGACCTCATTCAGACAGAAATCTATTACCTTTGCCATGCTCTGGTGAACA 490
DB 446 ATGATCCAGATGACCTCATTCAGACAGAAATCTATTACCTTTGCCATGCTCTGGTGAACA 505

QY 491 GAGAAATACACTGTTTATTTCTGAAATTTCCCAAACTATCAATAGAGCAGCAGCTTTAA 550
DB 506 GAGAAATACACTGTTTATTTCTGAAATTTCCCAAACTATCAATAGAGCAGCAGCTTTAA 565

QY 551 TGCTCTCTGGAGGCTCTTTTGGATCTTTTTCAGGCAACACTGGACTATGGATGATTT 610
DB 566 TGCTCTCTGGAGGCTCTTTTGGATCTTTTTCAGGCAACACTGGACTATGGATGATTT 625

QY 611 CTGAGAGAGAGAACTATTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 670
DB 626 CTGAGAGAGAGAACTATTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 685

QY 671 ACGATTATACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 730
DB 686 ACGATTATACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 745

QY 731 TAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 790
DB 746 TAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 805

QY 791 CTAGTTGTGCTCAACATACATGAGTGGATCCAAATTTATGCCCGCTGATCCAGAGCTGAGATTG 850
DB 806 CTAGTTGTGCTCAACATACATGAGTGGATCCAAATTTATGCCCGCTGATCCAGAGCTGAGATTG 865

QY 851 CTTTATATACATGCAACGATATTTGGATATCTAACATCGTAAACCAAGAGAGAGAGAGAG 910
```

```
DB 866 CTTTATATACATGCAACGATATTTGGATATCTAACATCGTAAACCAAGAGAGAGAGAGAGAC 925
QY 911 TCACCTTATGTGCACAAATGAGCTAGCCAAATGGAAGAGATCCAGATCAGCTGGAGTCG 970
DB 926 TCACCTTATGTGCACAAATGAGCTAGCCAAATGGAAGAGATCCAGATCAGCTGGAGTCG 985
QY 971 CTACCTTTGTTCTCCAAAGAGAAATTTGATAGATATTTCTGGCTATTTGGTGGTGTCCAAAG 1030
DB 986 CTACCTTTGTTCTCCAAAGAGAAATTTGATAGATATTTCTGGCTATTTGGTGGTGTCCAAAG 1045

QY 1031 CTGAAAACAACTCCACAGTGGTGTAAATTTCTTAGAATTTCTATATCAAGAAATATGATGAT 1090
DB 1046 CTGAAAACAACTCCACAGTGGTGTAAATTTCTTAGAATTTCTATATCAAGAAATATGATGAT 1105

QY 1091 CTGAGGTGGAAATTTATTCATGTTTACATCCCTTATGTTTACATCCCTTATGTTTACAT 1150
DB 1106 CTGAGGTGGAAATTTATTCATGTTTACATCCCTTATGTTTACATCCCTTATGTTTACAT 1165

QY 1151 TCCGTTATCTTAAACAGGTTACAGCAATCTTAAAGTCACCTTTTAAAGATGTCAGAAATAA 1210
DB 1166 TCCGTTATCTTAAACAGGTTACAGCAATCTTAAAGTCACCTTTTAAAGATGTCAGAAATAA 1225

QY 1211 TGATTGATGCTGAAGAGAGATCATAGATGTCATAGATAAGGAATTAATTCACCTTTTG 1270
DB 1226 TGATTGATGCTGAAGAGAGATCATAGATGTCATAGATAAGGAATTAATTCACCTTTTG 1285

QY 1271 AGATTCATTTGAGGAGTTGAATATATTTGCCAGAGCTGGATGGACTCCTCAGGAGAAAT 1330
DB 1286 AGATTCATTTGAGGAGTTGAATATATTTGCCAGAGCTGGATGGACTCCTCAGGAGAAAT 1345

QY 1331 ATGCTTGGTCCATCTCTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCAC 1390
DB 1346 ATGCTTGGTCCATCTCTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCAC 1405

QY 1391 CTGAATTTATTTATCCAGTAGAAGATGATGTTTATGGAAGGAGAGAGACTCAATCAGTCAG 1450
DB 1406 CTGAATTTATTTATCCAGTAGAAGATGATGTTTATGGAAGGAGAGAGACTCAATCAGTCAG 1465

QY 1451 TGCTGATTTCTGTGACGCCCACTAAATTTATCTATGAAGAAACAAACAGACATCTGGATAATA 1510
DB 1466 TGCTGATTTCTGTGACGCCCACTAAATTTATCTATGAAGAAACAAACAGACATCTGGATAATA 1525

QY 1511 TCCATGACATCTTTTCAATGTTTTCCTCCAAAGTCACGAAGAGGAAATGAGTTTATTTTG 1570
DB 1526 TCCATGACATCTTTTCAATGTTTTCCTCCAAAGTCACGAAGAGGAAATGAGTTTATTTTG 1585

QY 1571 CCTCTGAATGCAAAACAGGTTTCCGTCATTTATACAAATTTACATCTATTTTAAAGGAAA 1630
DB 1586 CCTCTGAATGCAAAACAGGTTTCCGTCATTTATACAAATTTACATCTATTTTAAAGGAAA 1645

QY 1631 GCAATATATAACGATCCAGTGGTGGCTGCTGCTCCAGAGTATTCAGTGTCTCTATCA 1690
DB 1646 GCAATATATAACGATCCAGTGGTGGCTGCTGCTCCAGAGTATTCAGTGTCTCTATCA 1705

QY 1691 AAGAGAGATPAGCAATTCACAGTGTGAAATGGGAAGTTCTTGGCCGCGATGGATCTTAATA 1750
DB 1706 AAGAGAGATPAGCAATTCACAGTGTGAAATGGGAAGTTCTTGGCCGCGATGGATCTTAATA 1765

QY 1751 TCCAAGTTGATGAAGTCAGAAGCTGGTATATTTTGAAGGCAACCAAGAGCTCCCTTTAG 1810
DB 1766 TCCAAGTTGATGAAGTCAGAAGCTGGTATATTTTGAAGGCAACCAAGAGCTCCCTTTAG 1825

QY 1811 AGCATCACCTGTACGTAGTACGTAAATCTCGGAGAGGTGACAGGCTGACTGACC 1870
DB 1826 AGCATCACCTGTACGTAGTACGTAAATCTCGGAGAGGTGACAGGCTGACTGACC 1885

QY 1871 GTGGCTACTCACATTTCTGCTGCATCAGTCAGCACTGTGACTTCTTTTATAGTAAATATA 1930
DB 1886 GTGGCTACTCACATTTCTGCTGCATCAGTCAGCACTGTGACTTCTTTTATAGTAAATATA 1945

QY 1931 GTAACAGAGAGAAATCCACTGTGTGCTCCCTTTTAAAGCTATCAAGCTATCAAGTCTGAGAGTACC 1990
```

Db	1946	GTAAACAGAGAAATCACACTGTGTGTCCTTTTACAGCTATCAAGTCCTGAAGATGACC	2005
Qy	1991	CAACTTGCACAAACAAAGGAATTTTGGGCCACACATTTTGGATTCAGCAGGTCCTCTTCCTG	2050
Db	2006	CAACTTGCACAAACAAAGGAATTTTGGGCCACACATTTTGGATTCAGCAGGTCCTCTTCCTG	2065
Qy	2051	ACTATACTCTCCAGAAATTTTCTTTTGAAGTACTACTGGATTTACATTCGTATGGGA	2110
Db	2066	ACTATACTCTCCAGAAATTTTCTTTTGAAGTACTACTGGATTTACATTCGTATGGGA	2125
Qy	2111	TGCTCTACAGCTCATGATCTACAGCCTGGAAGAAATATCCCTACTGTGCTGTCATAT	2170
Db	2126	TGCTCTACAGCTCATGATCTACAGCCTGGAAGAAATATCCCTACTGTGCTGTCATAT	2185
Qy	2171	ATGGTGGCTCCTCAGGTGCAATTTGGTGAATATCGGTTTAAAGGAGTCAAGTATTTCCGCT	2230
Db	2186	ATGGTGGCTCCTCAGGTGCAATTTGGTGAATATCGGTTTAAAGGAGTCAAGTATTTCCGCT	2245
Qy	2231	TGAATACCTAGCCTCTAGGTTATGTTGGTTAGTGTAGATAGACACAGGGGATCTCTGTC	2290
Db	2246	TGAATACCTAGCCTCTAGGTTATGTTGGTTAGTGTAGATAGACACAGGGGATCTCTGTC	2305
Qy	2291	ACCGAGGCTTAAATTTGAAGCGGCTTTAAATATAAAATGGTCAATAGAAATGACG	2350
Db	2306	ACCGAGGCTTAAATTTGAAGCGGCTTTAAATATAAAATGGTCAATAGAAATGACG	2365
Qy	2351	ATCAGTGTGAAGACTCCAAATATCTAGCTTCTCGATATGATTTCAATGACTTATGATCGTG	2410
Db	2366	ATCAGTGTGAAGACTCCAAATATCTAGCTTCTCGATATGATTTCAATGACTTATGATCGTG	2425
Qy	2411	TGGCATCCACGGCTGCTATGAGGATACCTCTCCCTGATGGCATTAATGACAGGTT	2470
Db	2426	TGGCATCCACGGCTGCTATGAGGATACCTCTCCCTGATGGCATTAATGACAGGTT	2485
Qy	2471	CAGATATCTTCAGGGTTGCTATTTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATA	2530
Db	2486	CAGATATCTTCAGGGTTGCTATTTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATA	2545
Qy	2531	CAGATACACGGAACGTTATATGAGTCACTTGTGGGGCCCCAGTCACTCTGTGGATCTTCTATG	2590
Db	2546	CAGATACACGGAACGTTATATGAGTCACTTGTGGGGCCCCAGTCACTCTGTGGATCTTCTATG	2605
Qy	2591	GATCTGGGCATGCAAGCAGAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCTTAC	2650
Db	2606	GATCTGGGCATGCAAGCAGAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCTTAC	2665
Qy	2651	ATGGTTTCTTGGATGAGATGTCATTTTGGCATACACAGTATATTAAGTATTTTGTAG	2710
Db	2666	ATGGTTTCTTGGATGAGATGTCATTTTGGCATACACAGTATATTAAGTATTTTGTAG	2725
Qy	2711	TGAGGCTGGAAGCCATATGATTTACAGATCTATCTCAGGAGACACAGCATAAGAG	2770
Db	2726	TGAGGCTGGAAGCCATATGATTTACAGATCTATCTCAGGAGACACAGCATAAGAG	2785
Qy	2771	TTCCTGAATCGGAGAACATTAATGAATCTGATCTTTTGGATCTTCAAGAAACCTTG	2830
Db	2786	TTCCTGAATCGGAGAACATTAATGAATCTGATCTTTTGGATCTTCAAGAAACCTTG	2845
Qy	2831	GATCAGTATTTGCTCTTAAAGTATATATTTTGGACCTGTGTAGAACTCTCTGTAT	2890
Db	2846	GATCAGTATTTGCTCTTAAAGTATATATTTTGGACCTGTGTAGAACTCTCTGTAT	2905
Qy	2891	ACATGGCTATTTAAACCAATGAGGAGTTTAAATCAACAGAAACACAGAAATTCATCATC	2950
Db	2906	ACATGGCTATTTAAACCAATGAGGAGTTTAAATCAACAGAAACACAGAAATTCATCATC	2965
Qy	2951	ACATTTTGAATCGGATGATTAACATCTACTCTCTGAAATAAATGTTGGTCCATGCGAGG	3010
Db	2966	ACATTTTGAATCGGATGATTAACATCTACTCTCTGAAATAAATGTTGGTCCATGCGAGG	3025
Qy	3011	GTCTACGGTTTGTGGTATTAATCTAATACCTTAACCCCACTGCTCAAAATCAATGATA	3070
Db	3026	GTCTACGGTTTGTGGTATTAATCTAATACCTTAACCCCACTGCTCAAAATCAATGATA	3085
Qy	3071	CATATTTCTGAGAGCCCAAGCAATACCAATAAGAAATTTACTAAAAA	3120
Db	3086	CATATTTCTGAGAGCCCAAGCAATACCAATAAGAAATTTACTAAAAA	3135
RESULT 2			
US-11-151-601-21			
; Sequence 21, Application US/11151601			
; Publication No. US20060003413A1			
; GENERAL INFORMATION:			
; APPLICANT: Millennium Pharmaceuticals, Inc.			
; APPLICANT: Meyers, Rachel E.			
; APPLICANT: Olandt, Peter J.			
; APPLICANT: Kapeller-Libermann, Rosana			
; APPLICANT: Curtis, Rory A. J.			
; APPLICANT: Williamson, Mark			
; APPLICANT: Weich, Nadine			
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASE, PHOSPHATASE,			
; TITLE OF INVENTION: AND PROTEASE FAMILY MEMBERS AND USES THEREOF			
; FILE REFERENCE: MP100-054PIRCPIONDIVIM			
; CURRENT APPLICATION NUMBER: US/11/151,601			
; CURRENT FILING DATE: 2005-06-13			
; PRIOR APPLICATION NUMBER: US 10/170,789			
; PRIOR FILING DATE: 2002-06-13			
; PRIOR APPLICATION NUMBER: US 09/797,039			
; PRIOR FILING DATE: 2001-02-28			
; PRIOR APPLICATION NUMBER: PCT/US01/06525			
; PRIOR FILING DATE: 2001-02-28			
; PRIOR APPLICATION NUMBER: US 60/186,061			
; PRIOR FILING DATE: 2000-02-29			
; PRIOR APPLICATION NUMBER: US 09/882,166			
; PRIOR FILING DATE: 2001-06-15			
; PRIOR APPLICATION NUMBER: PCT/US01/19269			
; PRIOR FILING DATE: 2001-06-15			
; PRIOR APPLICATION NUMBER: US 60/212,078			
; PRIOR FILING DATE: 2000-06-15			
; PRIOR APPLICATION NUMBER: US 09/934,406			
; PRIOR FILING DATE: 2001-08-21			
; PRIOR APPLICATION NUMBER: PCT/US01/26052			
; PRIOR FILING DATE: 2001-08-21			
; PRIOR APPLICATION NUMBER: US 60/226,740			
; PRIOR FILING DATE: 2000-08-21			
; Remaining Prior Application data removed - See File Wrapper or PALM.			
; NUMBER OF SEQ ID NOS: 45			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 21			
; LENGTH: 2649			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
US-11-151-601-21			
Query Match 81.6%; Score 2547; DB 17; Length 2649;			
Best Local Similarity 99.9%; Pred. No; 0;			
Matches 2647; Conservative 0; Mismatches 2; Indels 0; Gaps 0;			
Qy	214	ATGCGAGCAGCAATGGAACACAGAACAGCTGGGTGTGAGATATTTGAAACTGCGGACTGT	273
Db	1	ATGCGAGCAGCAATGGAACACAGAACAGCTGGGTGTGAGATATTTGAAACTGCGGACTGT	60
Qy	274	GAGGAGATATTAATCAGAGGATCGGCTTAAATTTGAGCCCTTTTATGTTGAGCGGTAT	333
Db	61	GAGGAGATATTAATCAGAGGATCGGCTTAAATTTGAGCCCTTTTATGTTGAGCGGTAT	120
Qy	334	TCCTGGAGTCAGCTTAAAAAGCTGTTGCGATACACAGAAATATATCATGGCTACATGATG	393
Db	121	TCCTGGAGTCAGCTTAAAAAGCTGTTGCGATACACAGAAATATATCATGGCTACATGATG	180
Qy	394	GCTAAGGCACCATGATTTTATGTTTGTGAAGAGAAATGATCCAGATGGACCTCATTTCA	453
Db	181	GCTAAGGCACCATGATTTTATGTTTGTGAAGAGAAATGATCCAGATGGACCTCATTTCA	240
Qy	454	GACAGAAATCTATTACCTTGGCCATGCTGGTGAGACAGAGAAATACACTGTTTATTTCT	513

Db 241 GACAGAACTATTACCTTGCCATGCTGGTGTGAGAACAGAGAAATACACTGTTTATCT 300
Qy 514 GAAATTCCTCCAAACTATCAATAGACGACGAGCTTAAATGCTCTCTGGAAGCCTCTTTTG 573
Db 301 GAAATTCCTCCAAACTATCAATAGACGACGAGCTTAAATGCTCTCTGGAAGCCTCTTTTG 360
Qy 574 GATCTTTTTCAGGCACACCTGGACTATGGAATGTAATCTCGAAGAAGAACAATATTAGA 633
Db 361 GATCTTTTTCAGGCACACCTGGACTATGGAATGTAATCTCGAAGAAGAACAATATTAGA 420
Qy 634 GAAAGAAACGCAATGGACAGCTCGGAATTCCTTACGATTAATCAACAAAGGAAGTGA 693
Db 421 GAAAGAAACGCAATGGACAGCTCGGAATTCCTTACGATTAATCAACAAAGGAAGTGA 480
Qy 694 ACATTTCTGTTTCAAGCCGGTAGTGAATTTATCACTGTAAGAAAGATGGAGGCCCAAGGA 753
Db 481 ACATTTCTGTTTCAAGCCGGTAGTGAATTTATCACTGTAAGAAAGATGGAGGCCCAAGGA 540
Qy 754 TTTACGCAACAACTTTAAGGCCCAATCTAGTGGAAACTAGTTGTCCCAACATACGGATG 813
Db 541 TTTACGCAACAACTTTAAGGCCCAATCTAGTGGAAACTAGTTGTCCCAACATACGGATG 600
Qy 814 GATCCAAATTTATGCCCCGCTGATCCAGACTGGATGCTTTTATACATAGCAACGATATT 873
Db 601 GATCCAAATTTATGCCCCGCTGATCCAGACTGGATGCTTTTATACATAGCAACGATATT 660
Qy 874 TGGATATCTAAACATCGTAACAGAGAAAGAGAGACTCACTTATGTGCACAATGAGCTA 933
Db 661 TGGATATCTAAACATCGTAACAGAGAAAGAGAGACTCACTTATGTGCACAATGAGCTA 720
Qy 934 GCCAACATGGAAGAGATCCAGATCAGCTGGAGTCCGCTACCTTTGTTCTCCAAAGAA 993
Db 721 GCCAACATGGAAGAGATCCAGATCAGCTGGAGTCCGCTACCTTTGTTCTCCAAAGAA 780
Qy 994 TTTGATAGATATTCTGGCTATTGTTGGTGTCCAAAGCTGAAACCTCCAGTGGTGT 1053
Db 781 TTTGATAGATATTCTGGCTATTGTTGGTGTCCAAAGCTGAAACCTCCAGTGGTGT 840
Qy 1054 ABAATTCCTAGAAATCTATCAAGAAATGATGAATCTGAGGTGGAATATTATCATGTT 1113
Db 841 ABAATTCCTAGAAATCTATCAAGAAATGATGAATCTGAGGTGGAATATTATCATGTT 900
Qy 1114 ACATCCCTATGTTGGAACAAAGAGGCGAGATTCACTCCGTTATCTCTAAACAGGTACA 1173
Db 901 ACATCCCTATGTTGGAACAAAGAGGCGAGATTCACTCCGTTATCTCTAAACAGGTACA 960
Qy 1174 GCAATCTCAAAGTCACTTTTAAGATGTCAGAAATAATGATGCTGGAAGGAAGGATC 1233
Db 961 GCAATCTCAAAGTCACTTTTAAGATGTCAGAAATAATGATGCTGGAAGGAAGGATC 1020
Qy 1234 ATAGATGTCATAGATAAGGAACATAATCAACCTTTTGAGATTTCTATTGGAAGGAGTTGAA 1293
Db 1021 ATAGATGTCATAGATAAGGAACATAATCAACCTTTTGAGATTTCTATTGGAAGGAGTTGAA 1080
Qy 1294 TATATTGCCAGAGCTGGATGGACTCTGAGGGAATAATGCTTGGTCCATCTCTACTAGT 1353
Db 1081 TATATTGCCAGAGCTGGATGGACTCTGAGGGAATAATGCTTGGTCCATCTCTACTAGT 1140
Qy 1354 CGCTCCAGACTCGCTCAGATAGTGTGATCTCACTGAATATTATATCCAGTAGNA 1413
Db 1141 CGCTCCAGACTCGCTCAGATAGTGTGATCTCACTGAATATTATATCCAGTAGNA 1200
Qy 1414 GATGATGTTATGGAAGGACAGACTCAATGAGTGGCTGATTTCTGTGAGCGCACTA 1473
Db 1201 GATGATGTTATGGAAGGACAGACTCAATGAGTGGCTGATTTCTGTGAGCGCACTA 1260
Qy 1474 ATTATCTATGGAAGGACAGACTCAATGAGTGGCTGATTTCTGTGAGCGCACTA 1533
Db 1261 ATTATCTATGGAAGGACAGACTCAATGAGTGGCTGATTTCTGTGAGCGCACTA 1320
Qy 1534 CCCCAGTACGAGAGGAAATTTGATTTATTTTGGCTCTGAAATGCAAAACAGGTTTC 1593

Db 1321 CCCAAAGTACGAGAGGAAATTTGAGTTTATTTTTGGCTCTGAAATGCAAAACAGGTTTC 1380
Qy 1594 CGTCATTTTATAAAATTTACATCTATTTTAAAGGAAACAAATATAAAACGATCCAGTGGT 1653
Db 1381 CGTCATTTTATAAAATTTACATCTATTTTAAAGGAAACAAATATAAAACGATCCAGTGGT 1440
Qy 1654 GGGCTGCTCTCCAAAGTGATTTCAAGTGTCTCTATCAAGAGAGAGATAGCAATTTACCAGT 1713
Db 1441 GGGCTGCTCTCCAAAGTGATTTCAAGTGTCTCTATCAAGAGAGAGATAGCAATTTACCAGT 1500
Qy 1714 GGTGAATGGGAAGTTCTTTGGCCGGCATGGATCTAATATCCAAAGTTGATGAAGTCAGAAGG 1773
Db 1501 GGTGAATGGGAAGTTCTTTGGCCGGCATGGATCTAATATCCAAAGTTGATGAAGTCAGAAGG 1560
Qy 1774 CTGTGATATTTTGAAGGACCAAGAGCTCCCTTTTAGAGCATCACTCTGACGTAGTCAGT 1833
Db 1561 CTGTGATATTTTGAAGGACCAAGAGCTCCCTTTTAGAGCATCACTCTGACGTAGTCAGT 1620
Qy 1834 TACGTAATCTCGAGAGGTGACAAAGGCTGACAGGCTGACCTGGCTACTCACATTTCTGTGC 1893
Db 1621 TACGTAATCTCGAGAGGTGACAAAGGCTGACAGGCTGACCTGGCTACTCACATTTCTGTGC 1680
Qy 1894 ATCAGTCAGCACGTGACTCTTCTTTATAGTAGTATAGTACCAAGAAATCCACACTGT 1953
Db 1681 ATCAGTCAGCACGTGACTCTTCTTTATAGTAGTATAGTACCAAGAAATCCACACTGT 1740
Qy 1954 GTGTCCCTTTTACAAAGCTATCAAGTCTGAAAGTGAACCAACTTGGCAAAAACAAAGGAATTT 2013
Db 1741 GTGTCCCTTTTACAAAGCTATCAAGTCTGAAAGTGAACCAACTTGGCAAAAACAAAGGAATTT 1800
Qy 2014 TGGGCCACCAATTTTGAATTCAGAGGTCTCTCTCTGACTATATCTCTCCAGAAAATTTTC 2073
Db 1801 TGGGCCACCAATTTTGAATTCAGAGGTCTCTCTCTGACTATATCTCTCCAGAAAATTTTC 1860
Qy 2074 TCTTTTGAAGTACTACTGATGATTTACATTTGATGAGGATGCTCTCAAGACCTCATGATCTA 2133
Db 1861 TCTTTTGAAGTACTACTGATGATTTACATTTGATGAGGATGCTCTCAAGACCTCATGATCTA 1920
Qy 2134 CAGCCTGGAAAGAAATATCTACTGTCTGTCTATATATGTTGTTGCTCTCAGAGTGCAGTTG 2193
Db 1921 CAGCCTGGAAAGAAATATCTACTGTCTGTCTATATATGTTGTTGCTCTCAGAGTGCAGTTG 1980
Qy 2194 GTGAATAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTTAGCTCTCTAGT 2253
Db 1981 GTGAATAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTTAGCTCTCTAGT 2040
Qy 2254 TATGTGTTGTAGTATAGACAAACAGGGGATCCCTGTCCACCGAGGCTTTAAATTTGAAGC 2313
Db 2041 TATGTGTTGTAGTATAGACAAACAGGGGATCCCTGTCCACCGAGGCTTTAAATTTGAAGC 2100
Qy 2314 GCCTTTAAATATAAAATGSGTCAATAGAAATTCACCATCAGGTGGAAGGACTCCATAT 2373
Db 2101 GCCTTTAAATATAAAATGSGTCAATAGAAATTCACCATCAGGTGGAAGGACTCCATAT 2160
Qy 2374 CTAGCTTCTCGATATGATTTCAATGACTTAGATCGTGGGCATCCACGGCTGCTCTAT 2433
Db 2161 CTAGCTTCTCGATATGATTTCAATGACTTAGATCGTGGGCATCCACGGCTGCTCTAT 2220
Qy 2434 GGAGGATACCTCTCCCTGATGGCAATTAATGAGAGTCAAGATCACTTTACAGGTTGCTATT 2493
Db 2221 GGAGGATACCTCTCCCTGATGGCAATTAATGAGAGTCAAGATCACTTTACAGGTTGCTATT 2280
Qy 2494 GCTGGGGCCCCAGTCACTCTCTGTGATCTTCTATGATACAGGATACAGGAACTTTATG 2553
Db 2281 GCTGGGGCCCCAGTCACTCTCTGTGATCTTCTATGATACAGGATACAGGAACTTTATG 2340
Qy 2554 GGTCAACCTTCAGCAAGAAATGAAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAA 2613
Db 2341 GGTCAACCTTCAGCAAGAAATGAAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAA 2400
Qy 2614 AAGTTCCCTCTGTAACCAAAATCGTTTACTGCTCTTACATCGTTTCTCTGGATGAGAATGTC 2673
Db 2401 AAGTTCCCTCTGTAACCAAAATCGTTTACTGCTCTTACATCGTTTCTCTGGATGAGAATGTC 2460

QY 2674 CATTTCACATACCACTATATTTACTAGTGTCTTTTGTAGTGGCTGGAAGCCATATGAT 2733
|||||
Db 2461 CATTTCACATACCACTATATTTACTAGTGTCTTTTGTAGTGGCTGGAAGCCATATGAT 2520
|||||
QY 2734 TTACAGATCTATCTCTCAGGAGACACAGCATAAGAGTTCTCTGAATCGGAGAACATTAT 2793
|||||
Db 2521 TTACAGATCTATCTCTCAGGAGACACAGCATAAGAGTTCTCTGAATCGGAGAACATTAT 2580
|||||
QY 2794 GAATGATCTTTTGGACTACTCTCAAGAAAACCTTGGATCAGTATTTGCTCTTAAAA 2853
|||||
Db 2581 GAATGATCTTTTGGACTACTCTCAAGAAAACCTTGGATCAGTATTTGCTCTTAAAA 2640
|||||
QY 2854 GTGATATAA 2862
|||||
Db 2641 GTGATATAA 2649
|||||

RESULT 3
US-09-925-065A-669313/c
; Sequence 669313, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 669313
; LENGTH: 1346
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-669313

Query Match 3.8%; Score 118; DB 7; Length 1346;
Best Local Similarity 100.0%; Pred. No. 3.4e-53;
Matches 118; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2036 CAGTCTCTCTCTGACTATCTCTCCAGAAATTTCTTTTGAAGTACTACTGGAT 2095
|||||
Db 118 CAGTCTCTCTCTGACTATCTCTCCAGAAATTTCTTTTGAAGTACTACTGGAT 59
|||||
QY 2096 TTACATTTGTATGGATGCTCTACAGCCTCATGATCTACAGCTCGAAAGAAATATCC 2153
|||||
Db 58 TTACATTTGTATGGATGCTCTACAGCCTCATGATCTACAGCTCGAAAGAAATATCC 1
|||||

RESULT 4
US-09-925-065A-743558/c
; Sequence 743558, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24

; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 743558
; LENGTH: 609
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-743558

Query Match 3.6%; Score 112; DB 7; Length 609;
Best Local Similarity 100.0%; Pred. No. 6.4e-50;
Matches 112; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2628 ACCAAATCGTTTACTGCTCTTACATGGTTTCTCGATGAGATGTCATTTTGCACATAC 2687
|||||
Db 199 ACCAAATCGTTTACTGCTCTTACATGGTTTCTCGATGAGATGTCATTTTGCACATAC 140
|||||
QY 2688 CAGTATATTACTGAGTTTTTTAGTGAGGCTGGAAGCCATATGATTACAG 2739
|||||
Db 139 CAGTATATTACTGAGTTTTTTAGTGAGGCTGGAAGCCATATGATTACAG 88
|||||

RESULT 5
US-09-925-065A-743559/c
; Sequence 743559, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 743559
; LENGTH: 609
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-743559

Query Match 3.6%; Score 112; DB 7; Length 609;
Best Local Similarity 100.0%; Pred. No. 6.4e-50;
Matches 112; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2628 ACCAAATCGTTTACTGCTCTTACATGGTTTCTCGATGAGATGTCATTTTGCACATAC 2687
|||||
Db 199 ACCAAATCGTTTACTGCTCTTACATGGTTTCTCGATGAGATGTCATTTTGCACATAC 140
|||||
QY 2688 CAGTATATTACTGAGTTTTTTAGTGAGGCTGGAAGCCATATGATTACAG 2739
|||||
Db 139 CAGTATATTACTGAGTTTTTTAGTGAGGCTGGAAGCCATATGATTACAG 88
|||||

RESULT 6
US-09-925-065A-818446/c

; CURRENT APPLICATION NUMBER: US/11/121,849
; CURRENT FILING DATE: 2005-05-03
; PRIOR APPLICATION NUMBER: 60/567,949
; PRIOR FILING DATE: 2004-05-03
; NUMBER OF SEQ ID NOS: 673904
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 207314
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapien
US-11-121-849-207314

Query Match 0.8%; Score 25; DB 17; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.033; 0; Indels 0; Gaps 0;
Matches 25; Conservative 0; Mismatches 0;

QY 2823 AAACCTTGGATCAGTATTGCTGCT 2847
|||||
Db 1 AAACCTTGGATCAGTATTGCTGCT 25

RESULT 11

US-11-121-849-207315
; Sequence 207315, Application US/11/121849
; Publication No. US20050272080A1
; GENERAL INFORMATION:

; APPLICANT: John Palma

; TITLE OF INVENTION: Methods of Genetic Analysis of Formalin Fixed Paraffin Embedded S

; FILE REFERENCE: 3684.1

; CURRENT APPLICATION NUMBER: US/11/121,849

; CURRENT FILING DATE: 2005-05-03

; PRIOR APPLICATION NUMBER: 60/567,949

; PRIOR FILING DATE: 2004-05-03

; NUMBER OF SEQ ID NOS: 673904

; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1

; SEQ ID NO 207315

; LENGTH: 25

; TYPE: DNA

; ORGANISM: Homo sapien

US-11-121-849-207315

Query Match 0.8%; Score 25; DB 17; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.033; 0; Indels 0; Gaps 0;
Matches 25; Conservative 0; Mismatches 0;

QY 2835 ACGTATTGCTGCTCTAAAGTGATA 2859
|||||
Db 1 ACGTATTGCTGCTCTAAAGTGATA 25

RESULT 12

US-11-121-849-207316
; Sequence 207316, Application US/11/121849
; Publication No. US20050272080A1
; GENERAL INFORMATION:

; APPLICANT: John Palma

; TITLE OF INVENTION: Methods of Genetic Analysis of Formalin Fixed Paraffin Embedded S

; FILE REFERENCE: 3684.1

; CURRENT APPLICATION NUMBER: US/11/121,849

; CURRENT FILING DATE: 2005-05-03

; PRIOR APPLICATION NUMBER: 60/567,949

; PRIOR FILING DATE: 2004-05-03

; NUMBER OF SEQ ID NOS: 673904

; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1

; SEQ ID NO 207316

; LENGTH: 25

; TYPE: DNA

; ORGANISM: Homo sapien

US-11-121-849-207316

Query Match 0.8%; Score 25; DB 17; Length 25;

Best Local Similarity 100.0%; Pred. No. 0.033;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2865 TTGACCTGTGTAGAACTCTCTGCTA 2889
|||||
Db 1 TTGACCTGTGTAGAACTCTCTGCTA 25

RESULT 13

US-11-121-849-207317
; Sequence 207317, Application US/11/121849
; Publication No. US20050272080A1
; GENERAL INFORMATION:

; APPLICANT: John Palma

; TITLE OF INVENTION: Methods of Genetic Analysis of Formalin Fixed Paraffin Embedded S

; FILE REFERENCE: 3684.1

; CURRENT APPLICATION NUMBER: US/11/121,849

; CURRENT FILING DATE: 2005-05-03

; PRIOR APPLICATION NUMBER: 60/567,949

; PRIOR FILING DATE: 2004-05-03

; NUMBER OF SEQ ID NOS: 673904

; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1

; SEQ ID NO 207317

; LENGTH: 25

; TYPE: DNA

; ORGANISM: Homo sapien

US-11-121-849-207317

Query Match 0.8%; Score 25; DB 17; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.033; 0; Indels 0; Gaps 0;
Matches 25; Conservative 0; Mismatches 0;

QY 2877 GAACTCTCTGTATACACTGGCTAT 2901
|||||
Db 1 GAACTCTCTGTATACACTGGCTAT 25

RESULT 14

US-11-121-849-207318
; Sequence 207318, Application US/11/121849
; Publication No. US20050272080A1
; GENERAL INFORMATION:

; APPLICANT: John Palma

; TITLE OF INVENTION: Methods of Genetic Analysis of Formalin Fixed Paraffin Embedded S

; FILE REFERENCE: 3684.1

; CURRENT APPLICATION NUMBER: US/11/121,849

; CURRENT FILING DATE: 2005-05-03

; PRIOR APPLICATION NUMBER: 60/567,949

; PRIOR FILING DATE: 2004-05-03

; NUMBER OF SEQ ID NOS: 673904

; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1

; SEQ ID NO 207318

; LENGTH: 25

; TYPE: DNA

; ORGANISM: Homo sapien

US-11-121-849-207318

Query Match 0.8%; Score 25; DB 17; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.033; 0; Indels 0; Gaps 0;
Matches 25; Conservative 0; Mismatches 0;

QY 2891 ACACCTGGCTATTAAACCAATGAGG 2915
|||||
Db 1 ACACCTGGCTATTAAACCAATGAGG 25

RESULT 15

US-11-121-849-207319
; Sequence 207319, Application US/11/121849
; Publication No. US20050272080A1
; GENERAL INFORMATION:

```

; APPLICANT: John Palma
; TITLE OF INVENTION: Methods of Genetic Analysis of Formalin Fixed Paraffin Embedded S
; FILE OF INVENTION: Microarrays
; FILE REFERENCE: 3684.1
; CURRENT APPLICATION NUMBER: US/11/121,849
; PRIOR FILING DATE: 2005-05-03
; PRIOR APPLICATION NUMBER: 60/567,949
; PRIOR FILING DATE: 2004-05-03
; NUMBER OF SEQ ID NOS: 673904
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 207319
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapien
; US-11-121-849-207319

Query Match      0.8%; Score 25; DB 17; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.033;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2930 GAAACACACAGAAATTCATCATCAT 2954
         |||||
Db       1 GAAACACACAGAAATTCATCATCAT 25

Search completed: May 4, 2006, 13:43:40
Job time : 1179 secs

```

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using **frame_plus_n2p** model

Run on: May 2, 2006, 03:09:14 ; Search time 90.4 Seconds
(without alignments)
3032.880 Million cell updates/sec

Title: US-10-825-632-2

Perfect score: 1030

Sequence: 1 aagtgttaagctccagg.....agaattactataaaaaaaa 3120

Scoring tables: **OLIGO**

Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 60.0 , Fgapext 7.0
Delop 60.0 , Delext 60.0

Searched: 2443163 seqs, 439378781 residues

Words **edges** 1

Total number of hits satisfying chosen parameters: 4589790

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODE=frame+n2p.model -DEV=xip
-Q=/abss/ABSSWEB_spool/US10825632/runat_01052006_110234_4411/app_query.fasta_1
-DB=A Geneseq -QFMT=fastan -SUFFIX=oligo.n2p.rag -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abs02p
-USER=US10825632@cgn_1_348_rsnat_01052006_110234_4411 -NCPU=6 -ICPU=3
-NO WMAP -NEG SCORES=0 -WAIT -DEBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=60 -FGAPEXT=7
-YGAPOP=60 -YGAPEXT=60 -DELOP=60 -DELEXT=60

Database : A_Geneseq_21.*

1: Geneseqp1980s.*
2: Geneseqp1990s.*
3: Geneseqp2000s.*
4: Geneseqp2001s.*
5: Geneseqp2002s.*
6: Geneseqp2003as.*
7: Geneseqp2003bs.*
8: Geneseqp2004s.*
9: Geneseqp2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	882	85.6	882	4 AAB47187	Aab47187 Human DPP
2	882	85.6	882	5 ABG61591	Abg61591 Human DPP
3	882	85.6	882	5 AAG78415	Aag78415 Amino aci
4	882	85.6	882	5 AAE24170	Aae24170 Human dip
5	882	85.6	882	5 AAU74749	Aau74749 Human pro
6	882	85.6	882	5 ADI17086	Adi17086 Human NOV
7	882	85.6	882	6 ABU07720	Abu07720 Human ser
8	706	68.5	746	8 ADI16334	Adi16334 Human pro
9	706	68.5	831	6 ABU92026	Abu92026 Human pro

10	700	68.0	824	6 ABU92030	Abu92030 Human pro
11	658	63.9	782	5 ABB97361	Abb97361 Novel hum
12	655	63.6	658	5 ABG61600	Abg61600 Human DPR
13	655	63.6	661	5 ABG61596	Abg61596 Human DPR
14	655	63.6	690	5 ABG61594	Abg61594 Human DPR
15	607	58.9	613	5 ABG61601	Abg61601 Human DPR
16	563	54.7	587	5 ADR41398	Adr41398 Human CD-
17	476	46.2	724	5 ABB97362	Abb97362 Novel hum
18	406	39.4	632	4 AAB93565	Aab93565 Human pro
19	341	32.1	465	4 AAB47189	Aab47189 Human DPP
20	339	32.9	358	5 ABG61597	Abg61597 Human DPR
21	275	26.7	738	8 ADT04044	Adt04044 Human pro
22	274	26.6	493	7 ADE78977	Ade78977 Human pro
23	262	25.4	360	4 AAB47190	Aab47190 Human DPP
24	250	24.3	250	5 ABB99949	Abb99949 Dipeptidyl
25	238	23.1	241	5 ABG61595	Abg61595 Human DPR
26	194	18.8	516	6 ABU92029	Abu92029 Human pro
27	194	18.8	580	5 AAE14337	Aae14337 Human pro
28	186	18.1	883	5 ADI17085	Adi17085 Murine NO
29	183	17.8	194	5 ABG61599	Abg61599 Human DPR
30	176	17.1	310	4 AAB47188	Aab47188 Human DPP
31	176	17.1	310	5 ABB08994	Abb08994 Human dip
32	176	17.1	310	7 ADD27858	Add27858 Human dip
33	151	14.7	193	5 ABB89739	Abb89739 Human pol
34	136	13.2	136	4 AAB74673	Aab74673 Human pro
35	86	8.3	108	5 ABG61598	Abg61598 Human DPR
36	79	7.7	129	4 AAM94906	Aam94906 Human rep
37	19	1.8	19	5 ABG61613	Abg61613 Human DPR
38	17	1.7	18	4 AAB47191	Aab47191 DPP8 pept
39	15	1.5	15	5 ABB99950	Abb99950 Dipeptidyl
40	15	1.5	16	4 AAB47192	Aab47192 DPP8 pept
41	13	1.3	432	6 AAB97157	Aab97157 Human dip
42	13	1.3	460	4 AAM38740	Aam38740 Human pol
43	13	1.3	512	4 AAM40526	Aam40526 Human pol
44	13	1.3	518	3 AAY90299	Aay90299 Human rep
45	13	1.3	518	5 ABG61603	Abg61603 Human DPR

ALIGNMENTS

RESULT 1

AAB47187

ID AAB47187 standard; protein; 882 AA.

AC AAB47187;

XX 29-JUN-2001 (first entry)

DT Human DPP8.

DE	Human DPP8.
XX	Human; dipeptidyl aminopeptidase; DPP8; prolyl oligopeptidase;
KW	dipeptidyl peptidase; DPPIV; T cell; cleavage; diarrhoea;
KW	growth hormone deficiency; glucose level; mucosal regeneration;
KW	non-insulin dependent diabetes mellitus; glucose intolerance;
KW	immunosuppression.
XX	Homo sapiens.
OS	Homo sapiens.
XX	Key Location/Qualifiers
FT	Active-site 739
FT	Active-site 817
FT	Active-site 849
FT	Active-site /note= "Forms part of Ser-Asp-His catalytic triad"
FT	Active-site /note= "Forms part of Ser-Asp-His catalytic triad"
FT	Active-site /note= "Forms part of Ser-Asp-His catalytic triad"
XX	WO200119866-A1.
XX	22-MAR-2001.
XX	11-SEP-2000; 2000WO-AU001085.
XX	10-SEP-1999; 99AU-00002762.

PR 18-FEB-2000; 2000AU-00005709.
XX (UNSY) UNIV SYDNEY.
XX PI Abbott CA, Gorell MD;
XX WPI; 2001-284520/29.
XX N-PSDB; AAC85694.
XX New human dipeptidyl aminopeptidase (DPP8) useful for cleaving
PT substrates, identifying inhibitors of DPP8 catalytic activity which have
PT therapeutic uses, and for detecting activated T cells.
XX Claim 1; Fig 2; 78pp; English.
XX This sequence represents human dipeptidyl aminopeptidase (DPP8). DPP8 has
CC substrate specificity for H-Ala-Pro-pNA, H-Gly-Pro-pNA and H-Arg-Pro-pNA.
CC Therefore, it is a prolyl oligopeptidase and a dipeptidyl peptidase,
CC because it is capable of hydrolysing the peptide bond C-terminal to Pro
CC in each of these compounds. DPP8 is homologous with human DPPIV. DPP8 is
CC useful for cleaving a substrate, and for detecting an activated T cell
CC which involves measuring the level of DPP8 gene expression in a T cell.
CC The level of DPP8 expression is detected by detecting the amount of DPP8
CC RNA in the cell. It is also useful for identifying a molecule capable of
CC inhibiting DPP8 cleavage of the substrate by DPP8. Molecules identified as
CC inhibiting DPP8 catalytic activity may be useful for treating diarrhoea,
CC growth hormone deficiency, lowering glucose levels in non-insulin
CC dependent diabetes mellitus and other disorders involving glucose
CC intolerance, enhancing mucosal regeneration and as immunosuppressants
XX Sequence 882 AA;
SQ
Alignment Scores:
Pred. No.: 0 Length: 882
Score: 882.00 Matches: 882
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 85.6% Indels: 0
DB: 4 Gaps: 0
US-10-825-632-2 (1-3120) x AAB47187 (1-882)
QY 214 ATGGCAGCAGCAATGGAAACAGACAGCTGGGTGTGAGATATTTGAAACTGCGGACTGT 273
Db 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
QY 274 GAGGAGAAATATTGAATCACAGGATCGGCCTAAATGGAGCCCTTTTATGTTGAGCGGTAT 333
Db 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
QY 334 TCCTGGAGTCAGCTTAAAGCTGCTTCCGATACAGAAATATCATGGCTACATCATG 393
Db 41 SerTrpSerGlnLeuLysLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
QY 394 GCTAAGGCACCATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGGACCTCATTTCA 453
Db 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
QY 454 GACAGATCTTATACCTGCCATGCTGTGTGAGACAGAGAATATACACTGTTTATCT 513
Db 81 AspArgIleTyrLysLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
QY 514 GAAATTCCTCCAAACTATCAATAGACGAGCTTAATGCTCTCTTGAAGCCCTCTTTTG 573
Db 101 GluIleProLysThrIleAsnArgAlaValLeuMetLeuSerTrpLysProLeuLeu 120
QY 574 GATCTTTTTCAGGCAACACTGGACTATGGAATGTATTCGAGAAGAGAACTATTAA 633
Db 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuLeuArg 140
QY 634 GAAAGAAAACCGCATGGAAACAGTCGGAATTCCTTACGATTTATCAACAGGAAGTGA 693
Db 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160

QY 694 ACATTTCTGTTTCAAGCCGGTAGTGAAATTTATCAGTAAAAAGATGGAGGCCCAAGA 753
Db 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
QY 754 TTTACGCAACAACTTTAAGCCCAATCTAGTGGAAACTAGTTGTCCTCCCAACATACGGATG 813
Db 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
QY 814 GATCCAAAATATTATGCCCCGCTGATCCAGACTGGATTGCTTTTATATACATAGCAACATATT 873
Db 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
QY 874 TGGATATCTTAACATCGTAACACAGAGAAGAAAGAGACTCACATTATGTGCACAATCAGCTA 933
Db 221 TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu 240
QY 934 GCCAACATGGAAGAAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAAAGAA 993
Db 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu 260
QY 994 TTTGATAGATATTCTGGCTATTGGTGTCTCCAAAAGCTGGAACAACTCCCAAGTGGTGT 1053
Db 261 PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly 280
QY 1054 AAAATCTTGAATTTCTATATGAAGAATGATGATCTGAGCTGGAATATTTCATGTT 1113
Db 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal 300
QY 1114 ACATCCCTATGTTGAAACAAGAGGGCAGATTTCATTCCTGTTATCTTAAACAGGTACA 1173
Db 301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr 320
QY 1174 GCAAACTCTAAAGTCATCTTTTAAGATGTGCAGAAATAATGATGTGTAAGGAGGATC 1233
Db 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
QY 1234 ATAGATCTCATAGATAAGAACTAATTCACCTTTTGAAGATTCTATTATGAAGGAGTTGAA 1293
Db 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
QY 1294 TATATTCCTCCAGAGCTGGATGGACTCCTGAGGAAAATATGCTGTGCTCATCTACTAGAT 1353
Db 361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
QY 1354 CGCTCCAGACTCGCCTACAGATAGTGTGATCTCACCTGAAATTTATTTATCCCAAGTAGAA 1413
Db 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
QY 1414 GATGATCTTATGAAAGGCAGAGACTCATTTGAGTCAGTGCCTGATTTCTGTGACGCCACTA 1473
Db 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
QY 1474 ATTATCTATCAAGAAACACACACATCTGGATAAATATCCATGACATCTTTTCATGTTT 1533
Db 421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
QY 1534 CCCCAAGTCCAGAAAGAAATTTGATTTATTTTTCCTGCTGATGCAAAACAGGTTTC 1593
Db 441 ProGlnSerHisGluGluGluIlePheIlePheAlaSerGluCysLysThrGlyPhe 460
QY 1594 CGTCATTTATACAAAATTACATCTATTTTTAAAGAAAGCAAAATATAACGATCCAGTGGT 1653
Db 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
QY 1654 GGGCTGCTCTCCCAAGTGAATTTCAAGTGTCTCTATCAAGAGGAGATAGCAATTTACAGT 1713
Db 481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluIleAlaIleThrSer 500
QY 1714 GGTGAATGGAAAGCTTTTGGCCGGCATGATCTAATATATCAAGTTGTATGAAGTCAGAGG 1773
Db 501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520

QY	1774	CTGGTATATTTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTACGTAGTCAGT	1833
Db	521	LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSer	540
QY	1834	TACCTAAATCCTGGAGAGGTGACAAGCGTCACTGACCGTGGCTACTACATTCCTGCTGC	1893
Db	541	TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys	560
QY	1894	ATCAGTCAGCACTGTGCTCTTTTATAGCTAGTAACTAGTAACTAGTAACTAGTAACTAGT	1953
Db	561	IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys	580
QY	1954	GTGTCCTCTTACAACTATCAAGTCTCTGAAGATGACCCAACTTGCAAAACAAAGAAATTT	2013
Db	591	ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe	600
QY	2014	TGGCCCAACATTTTGGATTGAGAGTCTCTCTCTCTGACTATCTCTCCAGAAATTTTC	2073
Db	601	TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIlePhe	620
QY	2074	TCTTTTCAAGTACTACTGATTTACATTTGATGGGATGCTCTCAAGCCTCATGATCTA	2133
Db	621	SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu	640
QY	2134	CAGCTCGAAGAAATATCTACTGTGCTGTTTATATATGTTGGTCTCTCAGGTGCACTG	2193
Db	641	GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu	660
QY	2194	GTGAATATCGTTTAAAGAGTCAAGTATTTCCCGTCTGAATACCTAGCTCTCTAGGT	2253
Db	661	ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly	680
QY	2254	TATCTGTTGTAGTAGACACACAGGGATCTGTCCACGAGGCTTTAAATTTGAAGGC	2313
Db	681	TyrValValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly	700
QY	2314	GCCTTTAAATATAAATGGTCAAAATAGAAATGACGATCAGGTGGAAGGACTCCAATAT	2373
Db	701	AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr	720
QY	2374	CTAGCTTCTCGATATGATTTTCATTTAGATCTGTGGGATCCAGGCTGCTCTAT	2433
Db	721	LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrSerTyr	740
QY	2434	GGAGTACCTCTCCCTGATGGCATTAAATGACAGGTACAGATCTCTCAGGTTGCTATT	2493
Db	741	GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle	760
QY	2494	GCTGGGCCCCAGTCACTCTGTGGATCTTCTATGATPACAGGATACACCGAAGCTTATATG	2553
Db	761	AlaGlyAlaProValThrLeuTyrPhePheTyrAspThrGlyTyrThrGluArgTyrMet	780
QY	2554	GTCACCTGACAGATGAACAGGGCTATTACTAGGATCTGTGGCCATCGACGACGAA	2613
Db	781	GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu	800
QY	2614	AAGTTTCCCTCTGAACCAATCGTTTACTGCTCTTACATGTTTCTCTGATGAGAATGTC	2673
Db	801	LysPheProSerGluProAsnArgLeuLeuLeuLeuHisGlyPheLeuAspGluAsnVal	820
QY	2674	CATTTTGCATACCATATATTACTAGATTTTATGAGGCTGGAAAGCCATATGAT	2733
Db	821	HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp	840
QY	2734	TTCAGATCTATCTCAGGAGACACACAGATGAAGTTCCTGAATCGGAGACATAT	2793
Db	841	LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr	860
QY	2794	GAATGCTATCTTTTGCATCTACCTTCAAGAAACCTTTGGATCAGTATTCGCTCTAAA	2853
Db	861	GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaLeuLys	880
QY	2854	GTGATA	2859

Db	881	ValIle	882
RESULT 2			
ABG61591			
ID	ABG61591	standard; protein; 882 AA.	
XX	XX		
AC	ABG61591;		
XX	XX		
DT	12-AUG-2002	(first entry)	
XX	XX		
DE	Human DPPIV related serine protease DPPP-1.		
XX	XX		
KW	KW	Human; serine protease; dipeptidyl peptidase IV-related protein; DPPP;	
KW	KW	DPPIV; infection; human immunodeficiency virus; HIV-1; HIV-2; pain;	
KW	KW	diabetes; infertility; obesity; anorexia; Parkinson's disease; cancer;	
KW	KW	heart failure; hypertension; urinary retention; osteoporosis; cancer;	
KW	KW	ulcer; allergy; cancer; psychotic disorder; neurological disorder;	
KW	KW	dyskinesia; reproductive disorder; inflammatory disorder;	
XX	XX	metabolic disorder.	
OS	OS	Homo sapiens.	
XX	XX		
PN	WO200231134-A2.		
PD	18-APR-2002.		
XX	XX		
PF	12-OCT-2001; 2001WO-US031874.		
XX	XX		
PR	12-OCT-2000; 2000US-0240117P.		
XX	XX		
PA	(FERR) FERRING BV.		
XX	XX		
PI	Qi S, Akinsanya KO, Riviere PJ, Junien J;		
XX	XX		
DR	WPI; 2002-444178/47.		
DR	N-PSDB; ABK83322.		
XX	XX		
PT	New dipeptidyl peptidase IV-related proteins and nucleic acids encoding		
PT	the proteins, useful for treating e.g: fungal, bacterial, protozoan and		
PT	viral infections, cancers, allergies, neurological disorders, or pain.		
XX	XX		
PS	Claim 17; Fig 1; 113pp; English.		
XX	XX		
CC	The present invention relates to the isolation of novel human serine		
CC	proteases referred to as dipeptidyl peptidase IV (DPPIV)-related proteins		
CC	(DPPP). The dipeptidyl peptidase IV-related proteins (DPPP) and nucleic		
CC	acids encoding them are useful for treating infections such as fungal		
CC	bacterial, protozoan and viral infections, particularly infections caused		
CC	by human immunodeficiency virus (HIV-1 or HIV-2), pain, diabetes,		
CC	precocious puberty, infertility, obesity, anorexia, bulimia, Parkinson's		
CC	disease, acute heart failure, hypotension, hypertension, urinary		
CC	retention, osteoporosis, angina pectoris, stroke, ulcers, asthma,		
CC	allergies, cancers, migraine, vomiting, psychotic and neurological		
CC	disorders (e.g. anxiety, dementia, or schizophrenia), and dyskinesias.		
CC	These may also be used in discovering therapeutic agents for the		
CC	treatment of reproductive, inflammatory and metabolic disorders. ABG61591		
CC	-ABG61612 represent human DPPP proteins		
XX	XX		
SQ	Sequence 882 AA;		
Alignment Scores:			
Pred. No.:	0	Length:	882
Score:	882.00	Matches:	882
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	85.6%	Indels:	0
DB:	5	Gaps:	0
US-10-825-632-2 (1-3120) x ABG61591 (1-882)			
QY	214	ATGCACGACGCAATGGAACAGACAGCTGGGTGTTGAGATATTTGAACACTCGGACTGT	273

Db 1 MetAlaAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaaspCys 20
QY 274 GAGGAGAAATATTGAATCACAGGATCGGCCTTAAATTCGAGCCCTTTTATGTGAGCGGTAT 333
Db 21 GluGluAenIleGluSerGlnaspArgProLysLeuGluProPheTyValGluArgTyr 40
QY 334 TCCTGGAGTCAGCTATAAAGCTGCTTGGCGATACAGAAATATCATGGCTACATGATG 393
Db 41 SerTrpSerGlnLeuLysLysLeuLeuAlaaspThrArgLysTyrHisGlyTyrMetMet 60
QY 394 GCTAAGGCACCATGATTTTCATGTTTGTGAGAGGAATGATCCAGATGGACCTCATCA 453
Db 61 AlalysAlaProHisaspPheMetPheValLysArgAsnaspProaspGlyProHisSer 80
QY 454 GACAGATCTATTACCTTGCATGCTGTGTGAGACAGAGAAAATACACTGTTTATTC 513
Db 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAenArgGluAenThrLeuPheTyrSer 100
QY 514 GAATTTCCAAACATCAATAGACGAGCAGTCTTAATGCTCTCTTGAAGCCCTCTTTTG 573
Db 101 GluIleProLysThrIleAenArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120
QY 574 GATCTTTTCAGCAACACTGGACTATGGAATGTATCTTCGAGAGAGAACTATTAGA 633
Db 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg 140
QY 634 GAAGAAGACGCATTTGGAACAGTCGGAATTTGCTTACGATTAATCAAGAGGAAGTGA 693
Db 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
QY 694 ACATTTCTGTTTCAAGCCGCTAGTGAATTTATACAGTAAAGATGGAGGCCACAGGA 753
Db 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
QY 754 TTTACGCAACACCTTTAAGGCCCAATCTAGTGGAACTAGTCTCCCAACATACCGATG 813
Db 181 PheThrGlnGlnProLeuArgProAenLeuValGluThrSerCysProAenIleArgMet 200
QY 814 GATCCAAATATTATGCCCTCGATCCAGACTGGATGCTGTTTATACATACACAGTATT 873
Db 201 AspProLysLeuCyProAlaaspProaspThrIleAlaPheIleHisSerAsnAspIle 220
QY 874 TGATATCTAACATCTACACGAGAGAAAGAGACTCACTTATGTGCAATAGCTA 933
Db 221 TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAenGluLeu 240
QY 934 GCAACATGGAGAGATCCAGATCAGCTGGATCGCTACCTTGTCTTCCAAAGAGAA 993
Db 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu 260
QY 994 TTTGATAGATATTCTGGCTATTGGTGTCCAAAGCTGAAACAACTCCCACTGGTGTG 1053
Db 261 PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly 280
QY 1054 AAAATCTTAGAATTTCTATGAGAAATGATGAATCTGAGGTGGAAATTTATTCATGTT 1113
Db 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal 300
QY 1114 ACATCCCTTATGTTGAAACAGGAGGCGAGATTCATTCGTTTATCCTAAACAGGTACA 1173
Db 301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr 320
QY 1174 GCAATCTCTAAAGTCACCTTTTAAAGATGTCAGAAATATGATTCATCTCAAGAGAGGATC 1233
Db 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
QY 1234 ATAGATGTCATAGATAAGGAATTAATTCACCTTTTGAGATTCATTTTCAAGAGGTTGAA 1293
Db 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
QY 1294 TATATTGCCAGAGCTGGAATGACCTCGAGGGAATATGCTTGGTCCATCTACTAGAT 1353
Db 361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380

QY 1354 CGCTCCAGACTCCCTTACAGATAGTGTGATCTCACCTGAATATTATATCCAGTAGAA 1413
Db 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
QY 1414 GATGATGTTATGGAAAGCGCAGAGACTCATTGAGTCAGTCCCTGATTTCTGTGACGCACTA 1473
Db 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
QY 1474 ATTATCTATGAAGAAACACACAGACATCTCGATAAATATCCATGACATCTTTCTATGTTTT 1533
Db 421 IleIleTyrGluGluThrThrAspIleTyrPheAsnIleHisAspIlePheHisValPhe 440
QY 1534 CCCCAAAGTCAACGAAGAGAAATTTGAGTTTATTTTGTCTCTGATGCAAAACAGGTTTC 1593
Db 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
QY 1594 CGTCATTTATACAAAATTTACATCTATTTTAAAGGAAGCAAAATATAACGATCCAGTGGT 1653
Db 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
QY 1654 GGGCTGCTGCTCCAAAGTGATTTCAAGTGTCTCTATCAAGAGGAGATAGCAATTTACCAGT 1713
Db 481 GlyLeuProAlaProSeraspPheLysCysProIleLysGluIleAlaIleThrSer 500
QY 1714 GGTCAATGGGAAGTCTTTCGCGGCAATGATCTATATCAAGTTGATGAAGTCAGAGG 1773
Db 501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArg 520
QY 1774 CTGGTATATTTGAAGGCACCAAGACTCCCTTTTAGACATCACTGTAGTAGTCAGT 1833
Db 521 LeuValTyrPheGluGlyThrLysaspSerProLeuGluHisLeuTyrValValSer 540
QY 1834 TAGCTAAATCTGGAGAGGTGCAAGGCTGACTGACCGTGGCTACTCACATTTCTGTGTC 1893
Db 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
QY 1894 ATCAGTCAGACTGTGACTCTTTTATAAGTATAGTAAACCAAGAGAATTCACACACTGT 1953
Db 561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
QY 1954 GTCTCCCTTTACAAGCTATCAAGTCTCGAAGATGACCAACTTGCAAAAACAAAGGAATTT 2013
Db 581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
QY 2014 TGGGCCACCATTTTGGATTCAGAGGTCTCTCTCTGACTATATCTCTCCAGAAATTTTC 2073
Db 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe 620
QY 2074 TCTTTTGAAGTACTACTGGATTTTACATTGTATGCGATGCTCTACAGCCTCATGATCTA 2133
Db 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
QY 2134 CAGCCTGGAAAGAAATATCTCTACTGTCTGTTCATATATGTTGGTCTCAGGTCCAGTTG 2193
Db 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu 660
QY 2194 GTCAATATCCGTTTAAAGGATCAAGTNTTCCGCTTGAATACCTTAGCTCTTAGGT 2253
Db 661 ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly 680
QY 2254 TATGTGGTGTAGTATAGACACAGGGATCTCTGCCAGGGCTTAAATTTTGAAGGC 2313
Db 681 TyrValValValIleAspAsnArgLysCysHisArgGlyLeuLysPheGluGly 700
QY 2314 GCCTTTAAATATAATGGGTCAAATGAAATTCAGATCAGGTGGAAGACTCCAAATAT 2373
Db 701 AlaPheLysTyrLysMetGlyGlnIleAspAspGlnValGluGlyLeuGlnTyr 720
QY 2374 CTAGTCTCTCGATATGATTTCAATTGACTTAGTTCGTGTGGCATCCACGGCTGGTCTTAT 2433
Db 721 LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyr 740

2434 GdAGGATACCTCCCTGATGGCATTAATGACAGAGTGCAGATATCTTCAGGGTGTCTATT 2493
2494 GCTGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACAGGATACGTTATATG 2553
2554 GGTCACTCCACAGCAATGACAGGCTATTACTTACGATCTGTGGCCATGACAGCAGAA 2613
2614 AAGTCCCTCTGACCAATCGTTTACTGCTCTTACATGTTCTTCTGATGAGAAATGTC 2673
2674 CATTTTGCACATACCACTATATTACTGAGTCTTTTGTAGTGGGCTGGAAGCCATATGAT 2733
2734 TTACAGATCTATCTCTCAGGAGACACAGCATAAGAGTTCTCTGAATCGGGAGAACATTAT 2793
2794 GAACGTGATCTTTTGCACTACCTTCAAGAAACCTTGGATCACGTTATGCTGCTTAAAA 2853
2854 GTGATA 2859
2881 ValIle 882

RESULT 3
AAG78415
ID AAG78415 standard; protein; 882 AA.
AC AAG78415;
DT 12-APR-2002 (first entry)
DE Amino acid sequence of 21953 human prollyl oligopeptidase.
KW 21953 prollyl oligopeptidase; antibody; proline; endopeptidase; cancer;
KW cardiovascular disease; autoimmune disease; atopic allergy;
KW neuronal disorder; vascular disorder; prostate disorder; cystostatic;
KW antidiabetic; antiarthritic; antiasthmatic; antiinflammatory;
KW diabetes mellitus; arthritis; multiple sclerosis; asthma;
KW Grave's disease; neuronal disorder; demyelinating disease.
XX Homo sapiens.
XX OS
XX PN
XX WO200179473-A2.
XX PD
XX 25-OCT-2001.
XX PF
XX 11-APR-2001; 2001WO-US040483.
XX PR
XX 18-APR-2000; 2000US-0197508P.
XX PA (MILL-) MILLENNIUM PHARM INC.
XX MEYERS RA, WILLIAMSON M;
XX PI
XX WPI; 2002-034353/04.
XX DR N-PSDB; AAH99934.
XX PT New polypeptides 21953, member of human prollyl oligopeptidase family,
XX useful as diagnostic targets and therapeutic agents for controlling
XX cancer, lymphoma and leukemia.
XX PS
XX Claim 1; Page 102-103; 12ipp; English.
XX This invention relates to an isolated 21953 human prollyl oligopeptidase.

CC Which is cytostatic, antidiabetic, antiarthritic, neuroprotective,
CC antihypertoid, dermatological, antipsoriatic, antiasthmatic,
CC ophthalmological, antiinflammatory, nootropic, antiparkinsonian,
CC anticonvulsant, gynaecological, vasotropic, antianginal, cardiant,
CC antiatherosclerotic, anorectic and metabolic in its action. Uses include
CC gene therapy, expression or activity of 21953 protein modulator, it is
CC useful for identifying a compound which binds to it and can be used in
CC preventing, treating or detecting a cellular proliferative or
CC differentiative disorder. The 21953 molecules can act as novel diagnostic
CC targets and therapeutic agents for controlling disorders associated with
CC the aberrant activity or degradation of peptide hormones e.g., disorders
CC associated with cell differentiation and proliferation such as cancer,
CC immune function, reproductive, neurological and cardiovascular function.
CC The 21953 molecules are thus useful for treating and preventing cellular
CC proliferative and differentiative disorders, haematopoietic neoplastic
CC disorders, immune disorders such as autoimmune diseases, diabetes
CC mellitus, arthritis, multiple sclerosis, asthma, Grave's disease,
CC neuronal disorders, demyelinating diseases, vascular disorders and
CC metabolism or pain disorders. This sequence represents the amino acid
CC sequence of 21953 human prollyl oligopeptidase
XX
SQ Sequence 882 AA;
Alignment Scores:
Pred. No.: 0 Length: 882
Score: 882.00 Matches: 882
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 85.6% Indels: 0
DB: 5 Gaps: 0
US-10-825-632-2 (1-3120) x AAG78415 (1-882)
QY 214 ATGGCAGCAGCAATGGAACACAGACAGCTGGTGTGAGATATTTGAAACGCGACTGT 273
DB 1 MetAlaAlaMetGluThrGluGluGluValGluLeuPheGluThrAlaAspCys 20
QY 274 GAGGAGAATATTGAATCAGAGATCGCGCTTAAATTTGGAGCCCTTTTATTTGAGCGPAT 333
DB 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyValGluArgTyr 40
QY 334 TCCTGGAGTCAGCTTAAAGCTGTTGCGATACAGAAATATCATCGGTACATGATG 393
DB 41 SerTrpSerGlnLeuLysLysLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
QY 394 GCTAAGGCACACATGATTTTCATGTTTGTGAAGAGATGATCCAGATGGACCTCATCA 453
DB 61 AlalysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
QY 454 GACAGATCTATTACCTTCCATGCTGCTGGTGAACAGAGAAATACACTGTTTATTCT 513
DB 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
QY 514 GAAATTTCCAAAACATCATAGCAGCAGGAGTCTTAAATGCTCTTGTGAAGCCTCTTTTG 573
DB 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120
QY 574 GATCTTTTTCAGGCACACACTGGACTATGGAATGTATTTCTCGAGAGAGAACTATTAGA 633
DB 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg 140
QY 634 GAAAGAAAACGCAATTGGAACAGTCGGAATGCTTCTTACGATATTCACCAAGGAAGTGA 693
DB 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
QY 694 ACATTTCTGTTTGAAGCCGCTAGTAGGAAATATATCAGTAAAGATGGAGGCCACAGA 753
DB 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
QY 754 TTTACGCAACACCTTTTAAGSCCAATCTAGTGGAACTAGTTGTCACCAACATGATG 813
DB 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200

DT 23-SEP-2002 (first entry)
XX Human dipeptidyl peptidase 8 (DPP8) protein.
XX Human; dipeptidyl peptidase; DPP; neoplasia; type II diabetes; cirrhosis;
KW autoimmunity; human immuno deficiency virus; HIV infection; cytostatic;
KW graft rejection; antidiabetic; antiinflammatory; immunosuppressive;
KW antitviral; enzyme.
XX Homo sapiens.
XX WO200234900-A1.
XX 02-MAY-2002.
XX 29-OCT-2001; 2001WO-AU001388.
XX 27-OCT-2000; 2000AU-00001078.
XX (UNSY) UNIV SYDNEY.
XX Abbott CA, Gorrell MD;
XX WPI: 2002-454646/48.
XX N-PSDB; RAD38956.
XX New dipeptidyl peptidase (DPP) peptides, useful for screening inhibitors
PT of DPP catalytic activity, which may be employed to treat e.g. neoplasia,
PT type II diabetes, cirrhosis, autoimmunity, graft rejection and HIV
PT infection.
XX Example; Fig 1; 91pp; English.
XX The present invention relates to dipeptidyl peptidase (DPP) proteins and
CC polynucleotides encoding such proteins. The DPP peptides are useful for
CC screening inhibitors of DPP catalytic activity. The inhibitors are useful
CC for treating neoplasia, type II diabetes, cirrhosis, autoimmunity, graft
CC rejection and HIV (human immuno deficiency virus) infection. The present
CC sequence is human DPP8 protein
XX Sequence 882 AA;
SQ
Alignment Scored:
Pred. No.: 0 Length: 882
Score: 882.00 Matches: 882
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 85.6% Indels: 0
DB: 5 Gaps: 0
US-10-825-632-2.(1-3120) x AAE24170 (1-882)
QY 214 ATGGCAGCAGCAATGGAACAGACAGCTGGGTGTTGAGATATTTGAAACTGCGAGCTGT 273
DB 1 MetAlaAlaAlaMetGluThrGluInLeuGlyValGluIlePheGluThrAlaAspCys 20
QY 274 GAGGAGATATTGAATCACAGGATCGGCCTAAATTTGGAGCTTTTATGTTGAGCGGTAT 333
DB 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
QY 334 TCCTGGAGTCAGCTTAAAGCTGCTTGGCCGATACAGAGAAATATCATGGCTACATGATG 393
DB 41 SerTrpSerGlnLeuLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
QY 394 GTTAAGGCCACATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGGACCTCATTTCA 453
DB 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
QY 454 GACAGATCTATTACCTGGCTGCTGTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 513
DB 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
QY 514 GAAATTTCCAAACATCATATAGACAGCAGCTCTTAATGCTCTCTTGGAGCCTTTTGG 573

Db 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120
QY 574 GATCTTTTTCAGGCAACACACTGGACTATGGAATGATATTCGAGAGAGAGAACTATTAGA 633
DB 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgIleGluGluLeuLeuArg 140
QY 634 GAAAGAAAACCATTTGGAACAGTCGGAATTCCTTACGATTATCACCAGGAAGTGA 693
DB 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
QY 694 ACATTTCTGTTTCAAGCCGCTAGTGGAAATTTATCAGCTAAAGATGGAGGGCCACAAGA 753
DB 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
QY 754 TTTAGCAACAACTTTAAGCCCAATCTAGTCGAAACCTAGTTGTCCTCCCAACATAGCATG 813
DB 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
QY 814 GATCCAAAATTTATGCCCGCTGATCCAGACTGGAATTCCTTTTATACATAGACACCATATT 873
DB 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
QY 874 TGGATATCTACATCGTAACACAGAGAGAAAGAGACTCCTATTATGTGCACAAATGAGCTA 933
DB 221 TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu 240
QY 934 GCCAACATGGAAGAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCCAAGAGAA 993
DB 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu 260
QY 994 TTTGATAGATATTCCTGGCTATTGCTGCTCCAAAAGCTGAAACAACTCCCAAGTGGTGT 1053
DB 261 PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly 280
QY 1054 AAAATTTCTAGAATTCCTATATGAAGAAATATGAAATCTGAGGTGGAATATTATCATGTT 1113
DB 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal 300
QY 1114 ACATCCCTATGTTGGAAACAGAGGGCGAGATTCATTCGTTTATCTCTTAAACAGGTACA 1173
DB 301 ThrSerProMetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThr 320
QY 1174 GCAAATCCTAAAGTCACCTTTTAAGATGTCAGAAATATGATTCATGCTCAGAGAGAGATC 1233
DB 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
QY 1234 ATAGATGTCATAGATAAGGAACCTAAATCAACCTTTTGGAGATTCCTATTTGAAGAGTGA 1293
DB 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
QY 1294 TATATTGTCAGAGCTGGATGACTCCTCGAGGAAAAATATGCTTGGTCCATCTCTACTAGAT 1353
DB 361 TyrIleAlaArgAlaGlyTyrThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
QY 1354 CGCTCCACAGATCGCCTACAGATAGTGTTCATCTACCTGGAATTTATATCCAGTAGAA 1413
DB 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
QY 1414 GATGATGTTATGGAAGCGCAGAGACTCATTTGAGTCAGTCGCTGATTCGTGACGCCACTA 1473
DB 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
QY 1474 ATTATCTATGAGAAACAACAGACATCTGATAAATATCCATGACATCTTTCATGCTTTTT 1533
DB 421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
QY 1534 CCCCAAGTACGAGAGAGGAAATTTGAGTTTATTTTTCCTCTGAAATGCAAAACAGGTTC 1593
DB 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
QY 1594 CQTCAATTTATCAAAATTTACATCTATTTTAAAGGAAGCAAAATATAAACAGTCCAGTGGT 1653

Db 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerIleTyrLysArgSerSerCly 480
QY 1654 GGGCTGCTCCCAAGTGATTTCAAGTGTCTTCAAGAGAGAGATAGCATTTACAGT 1713
Db 481 GlyLeuProAlaProSerAspPheLysGlyProIleLysGluGluIleAlaIleThrSer 500
QY 1714 GGTGAATGGGAGTCTTGGCCGCGCATGGATCTAATATCAAGTTGATGAGTCGAAGG 1773
Db 501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520
QY 1774 CTGGTATATTTGAAGGCAACAAGACTCCCTTTAGAGCATCACTGTACGTACGT 1833
Db 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisIleLeuTyrValValSer 540
QY 1834 TACGTAATCTCGAGAGGTGACAAGGCTGACTGACCGTGGCTACTACATTTCTGTGTC 1893
Db 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgLysTyrSerHisSerCysCys 560
QY 1894 ATCAGTCAGCACTGTGACTTCTTTATAGTAAGTAAGTATAGTAACAGAGAATCCACACTGT 1953
Db 561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
QY 1954 GTGTCCTTTTACAAGCTATCAAGTCTCGAAGATGACCCAACTTGCAAAACAAAGGAATTT 2013
Db 581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
QY 2014 TGGGCGACCATTTTGGATTCAGCAGGTCCTTCTTCGTACTATACCTCCAGAAATTTTC 2073
Db 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIlePhe 620
QY 2074 TCTTTTGAAGTACTACTGGATTTACATTTATGATGGATGCTCTCAAGCCTCATGATCTA 2133
Db 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
QY 2134 CAGCTCGAAGAAATATCTACTGTGCTTTCATATATGGTGGTCTCAGGTGCAGCTTG 2193
Db 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyProGlnValGlnLeu 660
QY 2194 GTGATATCGTTTAAAGAGTCAAGTATTTCCGCTGATATACCTAGCTCTCTAGGT 2253
Db 661 ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly 680
QY 2254 TATGTGTTGTAGTGATAGACAACAGGGATCTGTCCAGGAGGCTTAAATTTGAAGGC 2313
Db 681 TyrValValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly 700
QY 2314 GCCTTTAAATATAAATGGTCAAAATAGAAATGACGATCAGGTGGAAGGACTCCAATAT 2373
Db 701 AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr 720
QY 2374 CTAGCTTCTCGATATGATTTTCATTGACTTAGATCGTGTGGGCATCCAGGCTGCTCAT 2433
Db 721 LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyr 740
QY 2434 GGAGGATACCTCTCCCTGATGGCATTAATGCAGAGGTGAGATATCTTCAGGGTTGCTATT 2493
Db 741 GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760
QY 2494 GCTGGGCCCCAGTCATCTGTGGATCTTCTATGATACAGGATACACGGAACGTTATATG 2553
Db 761 AlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMet 780
QY 2554 GGTCACTCAGCAGAAATGAAGGCTATTACTTAGGATCTGTGGCATCAAGCAGAGAA 2613
Db 781 GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu 800
QY 2614 AAGTTCCCTCTCGAACCAATCGTTTACTGCTCTTACATGTTTCTCGATGGAATGTC 2673
Db 801 LysPheProSerGluProAsnArgLeuLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820
QY 2674 CATTTTCACATACCATGATATTTACTAGTTTTTTAGTGAGGCTGGAAAGCCATATGAT 2733
Db 821 HisPheAlaHisThrSerIleLeuLeuSerPheLeuValAlaGalaGlyLysProTyrAsp 840

QY 2734 TTACAGATCTATCTCTCAGGAGAGACACAGCATTAAGAGTTTCTGATCGGAGAACATTAT 2793
Db 841 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860
QY 2794 GAATCGCATCTTTTGGCACTACCTTCAAGAAAACCTTTGGATCAGTATGCTCTCTAAAA 2853
Db 861 GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaLeuLys 880
QY 2854 GTGATA 2859
Db 881 ValIle 882
RESULT 5
ID AAU74749
AAU74749 standard; protein; 882 AA.
XX AAU74749;
XX 09-APR-2002 (first entry)
XX Human protease PRTS-9 protein sequence.
XX Human; protease; PRTS; gastrointestinal; Crohn's disease; cancer;
KW cardiovascular; atherosclerosis; autoimmune disorder; dermatitis;
KW inflammatory disorder; acquired immunodeficiency syndrome; AIDS;
KW cell proliferative disorder; developmental disorder; epilepsy;
KW Duchenne muscular dystrophy; epithelial disorder; neurological disorder;
KW reproductive disorder; endometriosis.
XX Homo sapiens.
OS
XX
XX WO200198468-A2.
XX 27-DEC-2001.
XX 13-JUN-2001; 2001WO-US019178.
XX 16-JUN-2000; 2000US-0212336P.
PR 22-JUN-2000; 2000US-0213955P.
PR 29-JUN-2000; 2000US-0215396P.
PR 07-JUL-2000; 2000US-0216821P.
PR 14-JUL-2000; 2000US-0218946P.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX Yue H, Elliott VS, Gandhi AR, Lal P, Au-Young J, Tribouley CM;
PI Delegeane AM, Baughn MR, Nguyen DB, Lee EA, Hafalia A, Khan FA;
PI Walia NK, Yao MG, Lu DAM, Patterson C, Tang YT, Walsh RT;
PI Azimzai Y, Lu Y, Ramkumar J, Xu Y, Reddy R, Das D, Kearney L;
PI Kailick DA;
XX WPI; 2002-090437/12.
XX N-PSDB; ABK12892.
XX Twenty one human proteases (referred to as PRTS-1 to PRTS-21), useful in
PT the diagnosis, treatment and prevention of gastrointestinal (e.g.
PT gastritis), cardiovascular (e.g. atherosclerosis) and cell proliferative
XX (e.g. cancer) disorders.
PS Claim 1; Page 140-142; 177pp; English.
XX The present invention relates to twenty one new human proteases, referred
CC to as PRTS-1 to PRTS-21. The PRTS polynucleotides and polypeptides of the
CC invention are useful in the diagnosis, treatment and prevention of
CC gastrointestinal e.g. gastritis, esophageal carcinoma and Crohn's
CC disease, cardiovascular e.g. atherosclerosis, hypertension and myocardial
CC infarction, autoimmune/inflammatory e.g. acquired immunodeficiency
CC syndrome (AIDS), allergies and osteoarthritis, cell proliferative e.g.
CC cancer, developmental e.g. Duchenne and Becker muscular dystrophy,
CC epithelial e.g. dermatitis, neurological e.g. epilepsy and Alzheimer's
CC disease and reproductive e.g. infertility and endometriosis disorders.
CC Numerous other examples of each disorder are given in the specification.

The present protein sequence represents the human protease PRTS-9 protein of the invention									
CC									
XX									
SQ	Sequence 882 AA;								
Alignment Scores:									
Pred. No.:	0	Length:	882						
Score:	882.00	Matches:	882						
Percent Similarity:	100.0%	Conservative:	0						
Best Local Similarity:	100.0%	Mismatches:	0						
Query Match:	85.6%	Indels:	0						
DB:	5	Gaps:	0						
US-10-825-632-2 (1-3120) x AAU74749 (1-882)									
QY	214	ATGGCAGCAGCAATGGAACAGACAGCTGGCTGTTGAGATATTTGAAATCGCGACTGT	273						
Db	1	MetAlaAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys	20						
QY	274	GAGGAGAATATTGAATCAGCATCGGCCTAAATTTGGAGCCTTTTATGTTGAGCGGTAT	333						
Db	21	GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr	40						
QY	334	TCTCGAGTCAGCTTAAAAAGCTGCTTCCGATACCAAGAAATATCATGGCTACATGATG	393						
Db	41	SerTrpSerGlnLeuLysLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet	60						
QY	394	GCTAAGGCACACATGATTTTCATGTTGTGAAGAGGAATGATCCAGATGGACCTCATTC	453						
Db	61	AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer	80						
QY	454	GACGAATCTATTACCTTGCCTGCTGCTGAGAACAGAGAAATACACTGTTTATTCT	513						
Db	81	AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer	100						
QY	514	GAAATCCCCAAACTATCAATAGCAGCAGCTTAAATGCTCTCTTGAAGCCTCTTTTG	573						
Db	101	GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu	120						
QY	574	GATCTTTTTCAGGCAACACTGGACTATGGAATGTATTCTCGAGAAGAAGACTATTAGA	633						
Db	121	AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuLeuArg	140						
QY	634	GAAGAAGACGATTTGGACAGCTCGGAATGCTTCTTACGATTATCACCAAGGAGTGA	693						
Db	141	GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly	160						
QY	694	ACATTTCTGTTTCAGCCGGTAGTGGAAATTTATCACTAAAGATGGAGGGCCACAAGGA	753						
Db	161	ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly	180						
QY	754	TTTACGCAACACCTTTTAAGGCCCAATCTAGTGGAAACTAGTTGTCCCAACATACGGATG	813						
Db	181	PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet	200						
QY	814	GATCAAAATTTATGCCCGCTGATCCAGACTGGATTGCTTTTATACATAGCAACGATATT	873						
Db	201	AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle	220						
QY	874	TGGATATCTAACATCTGTAACAGAGAGAAAGAGACTCATTATGTGCACAATGAGCTA	933						
Db	221	TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu	240						
QY	934	GCACACATGGAAGACATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAGAAGAA	993						
Db	241	AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu	260						
QY	994	TTTGATAGATATTCTCGCTATTGGTGGTGTCACAAAGCTGAAACAACTCCCGATGGTGGT	1053						
Db	261	PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly	280						
QY	1054	AAAAATCTTAGAATTTCTATATGAAAGAAATGATCAATCTGAGGTGGAATTTATTCATGTT	1113						

281	LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleHisVal	300
1114	ACATCCCTCATGTTGGAAACAGGAGGCGAGATTCAATCCGCTTATCTCTAAACAGGTACA	1173
301	ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr	320
1174	GCAAAATCCTAAAGTCACATTTTAAAGATGTGAGAAATPAATGATGTGATGAGGAGGATC	1233
321	AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle	340
1234	ATAGATGTCATGATGAAGAACTAAATTCACCTTTTGGAGATCTTATTTGAAGGAGTTGAA	1293
341	IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu	360
1294	TATATTGCCAGACCTGGATGGACTCTCGAGGGAAATATGCTTGCTCATCTACTAGAT	1353
361	TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp	380
1354	CGTCCCGACACTCGCTACAGATAGTGTGATCTCCCTCGAATTTATTTATCCAGTAGAA	1413
381	ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu	400
1414	GATGATGTTATGAAAGCGAGACACTATTGATGTCAGTCGCCTGATTCGTGAGCGCCACT	1473
401	AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu	420
1474	ATTATCTATGAAGAAACAAGACATCTCGATGAATATATCCATGACATCTTTTCATGTTT	1533
421	IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe	440
1534	CCCCAAGTCCACAGAGGAAATGAGTTTATTTTGGCTCTCGAATGCAAAAACAGGTTTC	1593
441	ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe	460
1594	CGTCATTTATACAAAATTACATCTATTTTAAAGGAAAGCAAAATATAAACGATCCAGTGT	1653
461	ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly	480
1654	GGGCTGCTGCTCCAAGTGATTTCAAGTGTCCTATCAAAAGAGAGATACCAATTACCAGT	1713
481	GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer	500
1714	GGTGAATGGGAAGTCTTTGCGCGCATGGATCTAAATATCAAGTTGATCAAGTCAAGAG	1773
501	GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg	520
1774	CTGTATATTTTGAAGCCACAAAGACTCCCTTTAGAGCATCACCTGTACGTAGTCACT	1833
521	LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisIleLeuTyrValValSer	540
1834	TACGTAAATCCTCGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACATTTCTGCTGC	1893
541	TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys	560
1894	ATCAGTCAGCACTGTGACTTCTTTTAAGTAAGTATAGTAACCAAGAGAAATCCACTCT	1953
561	IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys	580
1954	GTGTCCTTTTCAAGCTTATCAAGTCTCAAGTCCGAAAGATGACCCCACTTGCAAAACAA	2013
581	ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe	600
2014	TGGGCCACCATTTTGGATTTCAGCAGGTCTCTTCTGACTATACTCTCCAGAAATTTTC	2073
601	TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe	620
2074	TCTTTTGAAGTACTACTGGATTTACATGTATGGTAGTCTTACAGCCCTCATGATCTA	2133
621	SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu	640
2134	CAGCCTGGAAGAAATATCTCTACTGCTGCTGTTCATATATGCTGGTCTCAGGTGCACTTG	2193
641	GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu	660

or physiological responses in a cell, tissue, organ or organism. Specifically, it refers to the use of biologically active fragments for diagnostic and prognostic assays and furthermore in the treatment of diverse pathological conditions. The present invention describes novel human and murine NOVX proteins, as well as methods to modulate their expression using antisense oligos, ribozymes and peptide nucleic acids. The NOVX polypeptides, polynucleotides and antibodies are useful in treating or preventing NOVX-associated disorders, e.g. cardiomyopathy, atherosclerosis, cancer and diabetes. Furthermore, they may be used in treating or preventing diseases such as inflammation, autoimmune disorders, allergies, blood disorders, acquired immunodeficiency syndrome (AIDS), obesity, asthma, immunoglobulin (Ig)A nephropathy, cirrhosis, arthritis, Alzheimer's disease, infections, stroke, muscular dystrophy and epilepsy. Accordingly, these molecules have many activities including cytoskeletal, anti-HIV, anti-diabetic, immunosuppressive, anti-allergic, haemostatic, anti-HIV, anti-diabetic, anti-arteriosclerotic, anorectic, antiasthmatic, nephrotropic, antiarthritic, hepatotropic, neuroprotective, nootropic, antibacterial, virucide, antiparasitic, relaxant and anticonvulsant. In addition, they are useful in screening assays to identify small molecules that modulate or inhibit, for example, neurogenesis, wound healing and angiogenesis. The nucleic acids are also used as in chromosome mapping, tissue typing, preventive medicine and pharmacogenomics. This polypeptide is a homologue of a human NOVX protein of the invention.

xx Sequence 882 AA;

Alignment Scores:
 Pred. No.: 0 Length: 882
 Score: 882.00 Matches: 882
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 85.6% Indels: 0
 DB: 5 Gaps: 0

US-10-825-632-2 (1-3120) x AD117086 (1-882)

214 ATGCGACGACGATGAAACAGACAGCTGGTGTGAGATTTTGAACCTCGGACTGT 273
 1 MetAlaAlaMetGluThrGluGluGluValGluLeuPheGluThrAlaAspCys 20
 274 GNGGAGATATTCGATCAGAGATCGCCTAAATTTGAGCGCTTTTATGTTGAGCGGTAT 333
 21 GluGluAsnLeuGluSerGluAspArgProLysLeuGluProPheTyrValGluArgTyr 40
 334 TCCTGGAGTCAGCTTAAAGCTGCTGCCGATACAGAAATATCATGCTCATCATG 393
 41 SerTyrSerGluLeuLysLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
 394 GCTAAGGACACATGATTTTCATGTTGTGAAGAGATGATCCAGATGGACCTCATTTCA 453
 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
 454 GACAGATCTATTACCTTGCATGCTGTGTGAGAACAGAGAAATATACACTGTTTATCT 513
 81 AspArgLysTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
 514 GAAATTCCTCAAACTATCAATAGACGACAGCTTAAATCTCTCTTGAAGACCTCTTTG 573
 101 GluLeuProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTyrLysProLeuLeu 120
 574 GATCTTTTTCAGGACACTGGACTATGGAATGATATTCGAGAGAGAACTATTAGA 633
 121 AspLeuPheGluAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuArg 140
 634 GAAAGAAACCATTTGGAACAGCTCGGAATTTGCTTCTTACCATTTATCAACAGAGAGTGA 693
 141 GluArgLysArgLysGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
 694 ACATTTCTGTTTCAAGCGGTAGTGAATTTATCATGTAAGATGAGGAGGCCACAGGA 753
 161 ThrPheLeuPheGluAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180

754 TTTTACGCAACACCTTTAAGCGCCCAATCTAGTGGAACTAGTTGTCCCAACATACGGATG 813
 181 PheThrGlnGluProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
 814 GATCCAAATTTATGCCCGCTGATCCAGATCGGATTTGCTTTTATACATAGCAACATATT 873
 201 AspProLysLeuCysProAlaAspProAspTyrIleAlaPheIleHisSerAsnAspIle 220
 874 TGGATATCTAATCTGTAACCCAGAGAAAGAGAGACTCATTTATGTGCACATAGCTTA 933
 221 TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu 240
 934 GCCAATCGAAGAGATGCCAGATCATCTGGAGTCGTACCTTTCTTCTCCAAAGAGAA 993
 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGluGlu 260
 994 TTTGTAGATATCTGGCTATTGTGTGTGTCGAAAGCTGAAACAACTCCAGTGTGTGT 1053
 261 PheAspArgTyrSerGlyTyrTyrProCysProLysAlaGluThrThrProSerGlyGly 280
 1054 AAAATTTCTAGAAATTTCTATATGAAGAAATGATGAATCTGAGTGGAAATTTATCATTT 1113
 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleHisVal 300
 1114 ACATCCCTATCTGGAACAAAGAGGCGCAGATTCATTCCTGTTATCTCTAAACAGGTACA 1173
 301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr 320
 1174 GCAATCCTAAAGTCACATTTTAAAGATGTCAGAAATATGATGTCGTGAAGGAAGATC 1233
 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
 1234 ATAGATGTCATAGATAAGGAACATAATCAACCTTTTGTGAGATTCATTTGAAGAGTTGAA 1293
 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
 1294 TATATTGCCAGAGCTGGATGGACTCTCAGGAAATATGCTTGGTCCATCTACTAGAT 1353
 361 TyrIleAlaArgAlaGlyTyrThrProGluGlyLysTyrAlaTyrSerIleLeuLeuAsp 380
 1354 CGCTCCAGACCTCGCTCAGATAGTGTGATCTCACCTGAAATTTATTTATCCCAAGTAGAA 1413
 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
 1414 GATGATGTTATGAAAGGAGAGACTCATTTGAGTCAGTGCCTGATTTCTGTGACCCACTA 1473
 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
 1474 ATTATCTGAAAGAAACACACACATCTGGATAAATATCCATCATCTTTTCATGTTTTT 1533
 421 IleIleTyrGluGluThrThrAspIleTyrPheIleAsnIleHisAspIlePheHisValPhe 440
 1534 CCCCAAGTCACGAAGGAAATTTGAGTTATTTTTGCTCTGAAATCAAAACAGGTTTC 1593
 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
 1594 CGTCATTTATACAAAATTTACATCTATTTTAAAGGAAACAAATATAACGATCCAGTGT 1653
 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
 1654 GGGCTGCTGCTCCAGTGAATTTCAAGTCTCCATCAAGAGAGAGATAGCAATTTACCAGT 1713
 481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer 500
 1714 GGTGAATGGGAAGTTCTTGGCCGCTGATCTTAATATCAAGTTGATGAAGTCAGAGG 1773
 501 GlyGluThrGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520
 1774 CTGTGTATATTTGAAGCACCAAGACTCCCTTTTGTAGCATCCTGTACGTACTAGT 1833
 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSer 540
 1834 TACGTAATCTCGAGAGGTGACAGGCTGACTGACCGTGTACTACATTTCTTGTCTGC 1893

Db 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
QY 1894 ATCAGTCAGCAGCTGTGAGCTCTTTTATAAGTAAGTATAGTAACACAGAGAAATCCACAGCTGT 1953
Db 561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
QY 1954 GTGTCCTCTTTAAAGCTATCAAGTCCTGAGAGATGACCCCACTTGCAGAACAGGAATTT 2013
Db 581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
QY 2014 TGGGCCACCATTTTGGATTCAGCAGCTCTCTCTCTGACTATCTCTCCAGAAATTTTC 2073
Db 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIlePhe 620
QY 2074 TCTTTTGAAGTACTACTGATTTACATTGTATGGGATGCTCTCAAGCCCTCATGATCTA 2133
Db 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
QY 2134 CAGCCTCGAAGAAATATCTACTGTCTGCTTCATATATATGGTGGTCTCAGGTGCAGTTG 2193
Db 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu 660
QY 2194 GTGAATATCGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCTAGCTCTCTAGGT 2253
Db 661 ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly 680
QY 2254 TATGTGTTGTAGTATAGACACACAGGGATCCTGTCCAGGGCTTAAATTTGAAGGC 2313
Db 681 TyrValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly 700
QY 2314 GCCTTTAAATPATAAATGGGTCAAATAGAAATTTGACGATCAGGTGGAGGACTCCAATAT 2373
Db 701 AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr 720
QY 2374 CTAGCTTCTCGATATGATTTGACTTAGATCGTGTGGCATCCACGGCTGCTCTAT 2433
Db 721 LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyr 740
QY 2434 GGAGGATACCTCTCCCTGATGGCATTAAATGACAGAGGTGAGATATCTTCAGGGTGTCTAT 2493
Db 741 GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760
QY 2494 GCTGGGGCCCCAGTCACTCTGTGGATCTTCTATCATGATACAGATACAGGAGGAGTATATG 2553
Db 761 AlaGlyAlaProValThrLeuThrPheTyrAspThrGlyTyrThrGluArgTyrMet 780
QY 2554 GGTCACTGACACAGAAATGAACAGGGCTATTACTTAGGATCTGTGGCCATGACAGAGAA 2613
Db 781 GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu 800
QY 2614 AAGTTCCTCTGACCAATCGTTTACTGCTCTTACATGGTTTCTGATGAGATGATGTC 2673
Db 801 LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820
QY 2674 CATTTTGCATACACAGTATATTACTGAGTTTCTGAGGGCTGGAAAGCCATATGAT 2733
Db 821 HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840
QY 2734 TTACAGATCTATCTCAGAGAGACACAGCATAGAGTTCCTGAATCGGGAGAACATTAT 2793
Db 841 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860
QY 2794 GNACTGATCTTTGCACTACCTTCAAGAAAACTTTGGATCATCGTATTGCTGCTCTAAAA 2853
Db 861 GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeuLys 880
QY 2854 GTGATA 2859
Db 881 ValIle 882

RESULT 7
ABU07720

ID XX ABU07720 standard; protein; 882 AA.
AC ABU07720;
XX 19-MAY-2003 (first entry)
XX Human serine protease HIPHUM46.
XX Human; enzyme; HIPHUM46; serine protease; gene therapy; osteoarthritis;
KW serine protease activity modulation; dipeptidyl peptidase activity;
KW musculoskeletal disease; Hepatitis B virus infection; myotonic dystrophy;
KW Alzheimer's disease; paraspranuclear palsy; Huntington's disease;
KW amyotrophic lateral sclerosis; malabsorption syndrome; lung disease;
KW irritable bowel syndrome; type I diabetes; faecal incontinence;
KW haemorrhoid; proctitis; rectal polyp; small bowel tumour; dyslexia;
KW colorectal tumour; ceroid lipofuscinosis; allergic encephalomyelitis;
KW multiple sclerosis.
OS Homo sapiens.
XX Key Location/Qualifiers
XX Region 259..260
FT Active-site 739 Paired glutamates of the beta propeller domain"
FT Active-site 817 /label= Catalytic_serine_residue
FT Active-site 849 /label= Catalytic_aspartate_residue
FT Active-site 849 /label= Catalytic_histidine_residue
XX GB2374869-A.
XX 30-OCT-2002.
XX 22-JAN-2002; 2002GB-00001404.
XX 23-JAN-2001; 2001GB-00001760.
XX (GLAX) GLAXO GROUP LTD.
XX Edbrooke MR, Lewis AP;
XX WPI; 2003-150703/15.
XX N-PSDB; ABX12255.
XX Identifying modulators of serine protease activity useful for treating
XX musculoskeletal diseases, by contacting cell expressing a novel serine
XX protease polypeptide with a compound and monitoring serine protease
XX activity.
XX Claim 10; Page 26-29; 38pp; English.
XX The invention relates to a method of identifying a substance that
XX modulates serine protease activity, comprising contacting a cell, such as
XX a neuronal cell, lung cell, intestinal cell or a cell infected with a
XX virus, expressing a serine protease polypeptide (HIPHUM 46), or its
XX variant having dipeptidyl peptidase activity, or a serine protease
XX isolated from the cell with a test substance, and monitoring for serine
XX protease activity. The method is useful for identifying a substance that
XX modulates serine protease activity. A modulator of the serine protease is
XX useful in the manufacture of a medicament for treatment or prophylaxis of
XX a musculoskeletal disease e.g. osteoarthritis, Hepatitis B virus
XX infection, Alzheimer's disease, Paraspranuclear palsy, myotonic
XX dystrophy, Huntington's disease, or amyotrophic lateral sclerosis.
XX Additional disease that may be created using modulators of the serine
XX protease include malabsorption syndromes, irritable bowel syndrome, lung
XX disease, type I diabetes, faecal incontinence, haemorrhoids, proctitis,
XX rectal polyps, small bowel tumours, colorectal tumours, anaemia,
XX dyslexia, ceroid lipofuscinosis, allergic encephalomyelitis, and multiple
XX sclerosis. The present sequence represents the amino acid sequence of the
XX human serine protease HIPHUM46
XX Sequence 882 AA;

Alignment Scores:

Pred. No.: 0 Length: 882
Score: 882.00 Matches: 882
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 85.6% Indels: 0
DB: 6 Gaps: 0

US-10-825-632-2 (1-3120) x ABU07720 (1-882)

QY 214 ATGCGAGCAATGGAACAGAACAGCTGGGTGTGAGATATTGAAACTGCGGACTGT 273
DB 1 MetAlaAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
QY 274 GAGGAGAAATATTGAATCACAGGATCGGCCTAAATTGGAGCCCTTTTATGTGACGGTAT 333
DB 21 GluGluAsnIleGluSerGlnAspA-gProLysLeuGluProPheTyr-ValGluArgTyr 40
QY 334 TCCTGGAGTCAGCTTAAAAAGCTGTGCCGATACAGAAAATATCATGGCTCATGATG 393
DB 41 SerTrpSerGlnLeuLysLeuLeuAlaAspThrArgLysTyr-HisGlyTyr-MetMet 60
QY 394 GCTAAGGACCAACATGATTTTCATGTTTGTGAAGAGAAATGATCCAGATGGACCTCATCA 453
DB 61 AlalysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
QY 454 GACAGAATCTATTACCTTGCATGCTGTGTGAGAGAGAAATACACTGTTTATTCT 513
DB 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyr-Ser 100
QY 514 GAAATTCCTCAAAACTATCAATAGACGACGAGTCTTAATGCTCTCTCTGTGAAAGCCCTTTTG 573
DB 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120
QY 574 GATCTTTTTCAGGCAACACATGGACTATGGAATGTATTCTCGAGAAAGAAACTATTAGA 633
DB 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuLeuArg 140
QY 634 GAAAGAAAACGATTTGAAACAGTCGGAAATGCTTCTTAGCATATACCAAGGAGTGA 693
DB 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyr-HisGlnGlySerGly 160
QY 694 ACATTTCTGTTTCAAGCCGTAGTGGAAATTTATCAGTAAAGATGAGGGGCCACAGGA 753
DB 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
QY 754 TTACGCAACACCTTTAAGCCCAATCTAGTGGAACTAGTTGTCCCAACATACGGATG 813
DB 181 PheThrGlnGlnProLeuA-gProAsnLeuValGluThrSerCysProAsnIleArgMet 200
QY 814 GATCCAAAATTTATCCCGCTGTATCCAGACTGGATGCTTTTATATACATAGCAACGATAT 873
DB 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
QY 874 TGGATATCTAATCATCGTAACAGAGAGAAGAGAGACTCATTTATGTGCACATGAGCTA 933
DB 221 TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyr-ValHisAsnGluLeu 240
QY 934 GCGCAACATGGAAGATGCGCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAGAAGAA 993
DB 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu 260
QY 994 TTGTAGATATTCTGCTATTGCTGTGTGTCGAAAGCTGAAACACTCCAGCTGGTGGT 1053
DB 261 PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly 280
QY 1054 AAAATTCCTAGAAATCTATATGAAGAAATGATGAATCTGAGGTGGAATATTATTCTGT 1113
DB 281 LysIleLeuA-gIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal 300
QY 1114 ACATCCCTATGTTGGAAACAAAGGAGGCGAGATTCCTCGTTATCTCTAAACAGGTACA 1173

DB 301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr 320
QY 1174 GCAATCTCTAAAGTCACTTTTAAAGATGTGAGAAATAATGATTGCTGGAAGGAGATC 1233
DB 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
QY 1234 ATAGATGCTCATAGATAAGGAACATAATCAACCTTTTGGAGATTTCTATTGGAAGGATTGAA 1293
DB 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
QY 1294 TATATTCAGAGCTGGATGGACTCTCTGAGGAAAAATATGCTGTGCTCATCTACTAGAT 1353
DB 361 TyrIleAlaAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
QY 1354 CGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGATATTATTTATCCAGTAGAA 1413
DB 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
QY 1414 GATCATCTTATGAAAAGGACAGACTCATTGAGTCAGTCCCTGATTTCTGTGACGCCACTA 1473
DB 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
QY 1474 ATTTATCTATCAAGAAAACACAGACATCTGGATAAATATCCATGACATCTTTCTGTTTTT 1533
DB 421 IleIleTyrGluGluThrThrAspIleTTrpIleAsnIleHisAspIlePheHisValPhe 440
QY 1534 CCCCAAGTCACGAAGGAAATGAGTTTATTTTTTGGCTCTGAAATGCAAAACAGGTTTC 1593
DB 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
QY 1594 CGTCATTTATCAAAAATTACATCTATTTTAAAGGAAGCAATATAAAGCATCCAGTCGT 1653
DB 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
QY 1654 GGGCTGCTCTCCAGTGAATTTCAAGTGTCTATCAAGAGAGAGATAGCAATATACCACT 1713
DB 481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer 500
QY 1714 GGTCAATGGGAAGTCTTTGGCCGCATGGATCTTAATATATCAAGTTGTAGTGAAGTCAGAGG 1773
DB 501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520
QY 1774 CTGGTATATTTTGAAGGCACCAAGACTCCCTTTTAGACATCACCTGTACGTAGTCACT 1833
DB 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSer 540
QY 1834 TACGTAAATCTTGAGAGGTGACAAGCTGACCTGACCGTGGCTACTCACATTTCTTGCTGC 1893
DB 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyr-SerHisSerCysCys 560
QY 1894 ATCAGTCAGCTGTGATCTTTTATAAGTAACTATAGTATAGTAAACAGAGAATCCACACTGT 1953
DB 561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
QY 1954 GTGTCCCTTTTCAAGCTATCAAGTCTCTGAGATGACCCAACTTGCAAAACAAAGGAATTT 2013
DB 581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
QY 2014 TGGGCCACCATTTTGGATTTCAGCAGGTGCTCTTCTCTGACTATATCTCTCCAGAAATTTTC 2073
DB 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe 620
QY 2074 TCTTTTGAAGTACTACTCGAATTTACATTTATGTGGATGCTCTTACAGCCTCATGATCTA 2133
DB 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
QY 2134 CAGCTCGAAAGAAATATCTCTACTGCTGTTTATATATATGCTGCTGCTCCTCAGTGCAGTTG 2193
DB 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyProGlnValGlnLeu 660
QY 2194 GTGAATATTCGGTTTAAAGGAGTCAGTATTTCCGCTTGAATACCTAGCCCTCTCTAGT 2253
DB 661 ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly 680

QY	634	GMAGAGAAACGCAATTTGGAACAGTCGGAATTCCTTCTTACGATTATCACCAAGAGAGTGA	693	Db	501	GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg	520
Db	141	GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly	160	QY	1774	CTGGTATATTTTGAAGGACCAAGACATCCCTTTTAGAGCATCACCTGTACGTAGTACGT	1833
QY	694	ACATTTCTGTTTCAAGCCGGTAGTGGATTTTACACGTAAAGATCGAGGGCCACACAGGA	753	Db	521	LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSer	540
Db	161	ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly	180	QY	1834	TACGTAAATCCTGGAGAGGTGACAGGTGACTGACCGTGGCTACTCACATTTCTTCTGTCG	1893
QY	754	TTTACCCCAACACCTTTAAGGCCCAATCTAGTGAATACTAGTTGTCCCAACATACGATG	813	Db	541	TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys	560
Db	181	PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet	200	QY	1894	ATCAGTCACACATCTGACTCTTTTATAGTAAGTATAGTAACCCAGAGAAATCCACACTGT	1953
QY	814	GATCCAAAATATATGCCCCCTCATCCAGACTGGATTGCTTTTATACATAGCAACATATT	873	Db	561	IleSerGlnHisCysAspPheIleSerLysTyrSerAsnGlnLysAsnProHisCys	580
Db	201	AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle	220	QY	1954	GTCTCCCTTTTACAAGCTATCAAGTCTCTGAAGTACCCCACTTGCACAAAACAAGGAATTT	2013
QY	874	TGGATATCTAAACATCGTAACAGAGAGAAAGAGACTACCTATGTGCACATGAGCTA	933	Db	581	ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe	600
Db	221	TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu	240	QY	2014	TGGGCCACCATTTTGGATTTCAGCAGGTCTCTTCTGACTATATCTCCTCAGGAAATTTTC	2073
QY	934	GCCAAATGGAAGAGATCCAGATCAGCTGAGTCGCTACCTTTGTTCTCAAGAGAA	993	Db	601	TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe	620
Db	241	AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu	260	QY	2074	TCCTTTGAAAGTACTACTGGATTACATTGTATGGATGCTCTACAAAGCCTCATGATCTA	2133
QY	994	TTTGTAGATATTTCTGCTATTGTTGGTGTCCAAAGCTGAAACACTCCCACTGGTGT	1053	Db	621	SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu	640
Db	261	PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly	280	QY	2134	CAGCCTGGAAAGAAATATCTACTGCTGCTGTTTATATATATATATGCTGCTCAGGTG	2193
QY	1054	AAATTTCTTGAATTTCTATATGAAGAAATGATGAATCTGAGGTGGAATTTATTCATGTT	1113	Db	641	GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu	660
Db	281	LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleHisVal	300	QY	2194	GTGAATAATCGGTTTAAAGAGCTCAAGTATTTCGCTTGAATACCTAGCTCTTAGGT	2253
QY	1114	ACATCCCTATGTTGAAACAAGAGGCGAGATTCATTCCTGTTTCTTAAACAAGGTACA	1173	Db	661	ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly	680
Db	301	ThrSerProMetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThr	320	QY	2254	TATGTGTTGTAGTATAGACAAAGGGAGTCTCTGACCGAGGCTTAAATTTGAAGGC	2313
QY	1174	GCAATCCTAAAGTCACTTTAAGATGTCAGAAATAATGATTGATGCTGAAGGAGGATC	1233	Db	681	TyrValValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly	700
Db	321	AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle	340	QY	2314	GCCTTTAAATATAAATG 2331	
QY	1234	ATGATGTCATAGTAAGAACTAATTCAACCTTTTGAATTTCTTATTTGAAGGAGTTGAA	1293	Db	701	AlaPheLysTyrLysMet 706	
Db	341	IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluValGlu	360	RESULT 9			
QY	1294	TATATTTCCAGAGCTGGATGGAATCCTCGAGGAAAATATGCTTGCTCATCTACTAGAT	1353	ABU92026			
Db	361	TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp	380	ID	ABU92026	standard; protein; 831 AA.	
QY	1354	CGTCCAGACTCGCTTACAGATAGTGTGATCTCACCTGAATTTATTTATCCAGTAGAA	1413	XX	ABU92026;		
Db	381	ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu	400	XX	15-JUL-2003	(first entry)	
QY	1414	GATCATGTTATGGAAGGAGAGACTCATTGAGTCAGTGCCTGATTCGTGACCCCACTA	1473	DT	Human	protein modification and maintenance molecule-6 (PMM-6).	
Db	401	AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu	420	XX	Human;	protein modification and maintenance molecule; PMM; cancer;	
QY	1474	ATTATCTATGAAGAAACAAGACATCTGATTAATATCCATGACATCTTTTCATGTTTTT	1533	KW	Human;	cell proliferation disorder; atherosclerosis; neurological disorder;	
Db	421	IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe	440	KW	epilepsy; Huntington's disease; stroke; immune disorder; allergy;		
QY	1534	CCCCAAGTCAAGAGGAGAAATTTGATTTTATTTTTCCTCTGAATGCAAAACAGGTTTC	1593	KW	inflammatory disorder; AIDS; developmental disorder; hypothyroidism;		
Db	441	ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe	460	KW	Cushing's syndrome; gastrointestinal disorder; epithelial disorder;		
QY	1594	CGTCATTTTATCAAAATTTACATCTATTTTTAAAGGAAAGCAATATAACGATCCAGTGT	1653	KW	infection; cytostatic; antiarteriosclerotic; anticonvulsant; nootropic;		
Db	461	ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly	480	KW	neuroprotective; cerebroprotective; anti-HIV; antiallergic; vulnary;		
QY	1654	GGGTGCTGCTCCAGTATTTCAAGTGTCTTCAAGAGGAGATAGCAATACCAAGT	1713	XX	antiinflammatory; thymomimetic.		
Db	481	GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer	500	OS	Homo sapiens.		
QY	1714	GGTGAATGGGAAGTTCTTGGCCGCGCATGGATCTAATCAAGTTGATGAAGTCAGAGG	1773	XX	WO2003031939-A2.		

PR 28-NOV-2001; 2001US-0334145P.
PR 06-DEC-2001; 2001US-0337451P.
PR 14-DEC-2001; 2001US-0340584P.
XX (INCY-) INCYTE GENOMICS INC.
XX Ramkumar J, Gorvad AE, Baughn MR, Emerling BM, Yang J, Lee SX,
PI Tran UK, Becha SD, Duggan BM, Lee EA, Griffin JA, Li JX;
PI Sprague WW, Hatalla AJ, Chawla NK, Lehr-Wason BW, Kable AE, Yue H;
PI Marquis JP, Yao MG, Richardson TW, Tang Y, Jin P, Chien D;
PI Bhatia U, Burrill JD, Lee S, Blake JJ, Ho A, Zheng W;
XX WPI: 2003-430274/40.
DR N-PSDB; ACA92421.
XX New human protein modification and maintenance molecules (PMM), useful
PT for diagnosing, treating and preventing diseases or conditions associated
PT with the aberrant PMM expression e.g. cancer, atherosclerosis, or
PT infections.
XX Claim 1; Page 242-243; 31pp; English.
XX The present invention relates to the isolation of human protein
CC modification and maintenance molecules (PMM), and the polynucleotide
CC sequences encoding them. A total of 40 PMM polypeptides (designated PMM
CC -1 to PMM-40) are disclosed. The sequences of the invention are useful
CC for diagnosing a condition or disease associated with the expression of
CC PMM in a subject, preparing a polyclonal or monoclonal antibody, and
CC generating an expression profile of a sample containing the
CC polynucleotides. The diseases or conditions associated with decreased
CC expression or overexpression of PMM are cell proliferation disorders
CC (e.g. cancer, atherosclerosis), neurological disorders (e.g. epilepsy,
CC Huntington's disease, stroke), immune/inflammatory disorders (e.g. AIDS,
CC allergies), developmental disorders (e.g. hypothyroidism, Cushing's
CC syndrome), gastrointestinal or epithelial disorders and infections. The
CC PMM polypeptides or their fragments are useful in screening compounds
CC for effecting diseases as agonists or antagonists of the polypeptides, or in
CC altering the expression of the target polynucleotide and compounds that
CC specifically bind to, or modulate the activity of the polypeptide.
CC ABU92021-ABU92060 represent the human PMM polypeptides of the invention
XX
SQ Sequence 831 AA;
Alignment Scores:
Pred. No.: 0 Length: 831
Score: 706.00 Matches: 706
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 68.5% Indels: 0
DB: Gaps: 0
US-10-825-632-2 (1-3120) x ABU92026 (1-831)
QY 214 ATGGCAGCAGCAATGGAACAGAACAGCTGGGTGTGAGATATTTGAAACTCGGACTGT 273
Db 1 MetAlaAlaMetGluThrGluGluLeuGlyValGluIlePheGluThrAlaAspCys 20
QY 274 GAGGAGAATATTGAATACAGGATCGGCTAAATTTGGAGCTTTTATGTTGAGCGGTAT 333
Db 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyValGluArgTyr 40
QY 334 TCCTGGAGTCAGCTTAAAGAGCTGTTGCCGATACAGAAATATCATGGCTACATGATG 393
Db 41 SerTyrSerGlnLeuLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
QY 394 GCTAAGGCACACATGATTTTCATGTTGTGAAGAGAAATGATCCAGATGGACCTCATCA 453
Db 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
QY 454 GACAGAACTTATACCTGGCATGCTGGTGAGACAGAGAAATACACTGTTTATTTCT 513
Db 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrPheTyrSer 100

QY 514 GAAATTCCTCCAAAACTATCAATAGACGACGAGCTTAAATGCTCTCTTGGAAAGCCTCTTTTG 573
Db 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120
QY 574 GATCTTTTTCAGGCACACTGGACTATGGAATGTATTCGAGAGAAGAAGAACTATTAGA 633
Db 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuLeuArg 140
QY 634 GAAAGAAAACGCAATGGAAACAGTCGGAATTTGCTTACGATTATCACCAGGAGGATGGA 693
Db 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
QY 694 ACATTTCTGTTTCAAGCCGCTAGTGAATTTATACAGTAAAGATGGAGGCCACAGGA 753
Db 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
QY 754 TTTACGCAACAACTTTAAGGCCCAATCTAGTGGAACTAGTGTCTCCCAATACGAGATG 813
Db 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
QY 814 GATCCAAAATATGCGCGCTGATCCAGACTGGATGCTTTTATACATAGCAACGATATT 873
Db 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
QY 874 TGGATATCTAACHTGTACACAGAGAAGAAAGGAGACTCACTTATGTGCACAAATGAGCTA 933
Db 221 TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu 240
QY 934 GCCAATCATGGAAGAAGATGCCAGATCATCGTGGAGTCGCTACCTTTGTCTCCAAGAGAA 993
Db 241 AlaAsnMetGluGluLysAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu 260
QY 994 TTTGATAGATATCTGGCTATTGGTGTGTCCTCAAAAGCTGAAACAACTCCACAGTGGTGT 1053
Db 261 PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly 280
QY 1054 AAAATCTTAGAATCTTATATGAAGAAATCATCAATCTGAGGTGGAAATATTTCATGTT 1113
Db 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal 300
QY 1114 ACATCCCTTATGTTGAAACCAAGGAGGCGAGATTCATTCGTTTATCTCTAAACAGGTACA 1173
Db 301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr 320
QY 1174 GCAAAATCTTAAAGTCACCTTTTAAGATGTCAGAAATATGATTCATGCTGAGGAAGATC 1233
Db 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
QY 1234 ATAGATGTCATAGATAAGGAACCTAATTCACCTTTTGAGATTCTTATTCGAGGAGTTGAA 1293
Db 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
QY 1294 TATATTGCCAGAGCTGGATGGACTCTCAGGAAATATGCTTGGTCCATCTCTACTAGAT 1353
Db 361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
QY 1354 CGCTCCAGACTCGCTACAGATAGTGTGATCTCACTCGAATATTATTTATCCAGTAGAA 1413
Db 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
QY 1414 GATGATGTTATGAAAAGCAGAGACTCATGAGTCAGTCGCTGATTTCTGTGACGCCACTA 1473
Db 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
QY 1474 ATTATCTATGAAGAAAACAAAGACATCTGGATAAATATCCATGACATCTTTTCATGTTT 1533
Db 421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
QY 1534 CCCCAGATCAGGAAGGAGAAATTTGATTTTATTTTTCCTCTGAAATGCAAAACAGGTTTC 1593
Db 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
QY 1594 CGTCATTTATACAAATTTACATCTATTTTAAAGGAAAGCAAAATATAAAGCATCCAGTGGT 1653

```
Db 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
QY 1654 GGGTGGCTGCTCAAGTGATTTCAAGTGCTCTATCAAGAGAGATAGCAATACCAAGT 1713
Db 481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer 500
QY 1714 GGTGAATGGGAAGTTCTTGGCCGCATGATCTTAATATATCAAGTTGATGAAGTCAGAGG 1773
Db 501 GlyGluTrpGluValLeuGlyA-gHisGlySerAsnIleGlnValAspGluValArgArg 520
QY 1774 CTGGTATATTTTGAAGGACAAAGACTCCCTTTAGACATCACCTGTACTGACTCAGT 1833
Db 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSer 540
QY 1834 TACGTAATCTCGAGAGGTGACAGGCTGACAGGCTGACGCTGCTACTCATCTTCTGCTGC 1893
Db 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
QY 1894 ATCAGTGCAGCACTGTGACTCTCTTTAAGTAAGTATAGTAACAGAGAAATCCACACTGT 1953
Db 561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
QY 1954 GTGTCCCTTTTCAAGCTATCAAGTCTGAAGATGACCCAACTTGGCAACAAGCAATTT 2013
Db 581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
QY 2014 TGGGCCACATTTTGGATTCAGCAGGCTCTCTCTGACTATATCTCTCCAGAAATTTTC 2073
Db 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspPyrThrProGluLulePhe 620
QY 2074 TCTTTTGAAGTACTTGGATTTACATTTGATGGGATGCTCTACAGGCTCATGATCTA 2133
Db 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
QY 2134 CAGCTTGGAAAGAAATATCTACTGTGCTGTTCTATATATGTTGCTCTCAGTGCAAGTG 2193
Db 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu 660
QY 2194 GTGAATACGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTAGCTCTCTAGT 2253
Db 661 ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly 680
QY 2254 TATGCTGGTGTAGTGATACACACAGGAGGATCCCTGTCCAGGAGGCTTAAATTTGAAGC 2313
Db 681 TyrValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly 700
QY 2314 GCCTTTAAATATAAATG 2331
Db 701 AlaPheLysTyrLysMet 706
RESULT 10
ABU92030
ID ABU92030 standard; protein; 824 AA.
XX AC ABU92030;
XX DT 15-JUL-2003 (first entry)
XX DE Human protein modification and maintenance molecule-10 (PMM-10).
XX KW Human; protein modification and maintenance molecule; PMM; cancer;
KW cell proliferation disorder; atherosclerosis; neurological disorder;
KW epilepsy; Huntington's disease; stroke; immune disorder; allergy;
KW inflammatory disorder; AIDS; developmental disorder; hypothyroidism;
KW Cushing's syndrome; gastrointestinal disorder; epithelial disorder;
KW infection; cytostatic; antiarteriosclerotic; anticonvulsant; nootropic;
KW neuroprotective; cerebroprotective; anti-HIV; antiallergic; vulnerary;
KW antiinflammatory; thyromimetic.
XX OS Homo sapiens.
XX PN W02003031939-A2.
```

```
XX PD 17-APR-2003.
XX PF 11-OCT-2002; 2002WO-US032850.
XX PR 12-OCT-2001; 2001US-0329689P.
XX PR 25-OCT-2001; 2001US-0335703P.
XX PR 09-NOV-2001; 2001US-0348887P.
XX PR 28-NOV-2001; 2001US-0334145P.
XX PR 06-DEC-2001; 2001US-0337451P.
XX PR 14-DEC-2001; 2001US-0340584P.
XX PA (INCYTE GENOMICS INC.
XX PI Ramkumar J, Gorvad AE, Baughn MR, Emerling BM, Yang J, Lee SY;
XX PI Tran UK, Becha SD, Duggan BW, Lee EA, Griffin JA, Li JX;
XX PI Sprague WP, Hatalia AJA, Chawla NK, Lehr-Mason PM, Kable AE, Yue H;
XX PI Marquis JJ, Yao MG, Richardson TW, Tang TY, Jin P, Chien D;
XX PI Bhatia U, Burrill JD, Lee S, Blake JJ, Ho A, Zheng W;
XX PI WPI; 2003-430274/40.
XX DR N-PSDB; ACA92425.
XX PT New human protein modification and maintenance molecules (PMM), useful
XX PT for diagnosing, treating and preventing diseases or conditions associated
XX PT with the aberrant PMM expression e.g. cancer, atherosclerosis, or
XX PT infections.
XX PS Claim 1; Page 249-251; 31pp; English.
XX CC The present invention relates to the isolation of human protein
XX CC modification and maintenance molecules (PMM), and the polynucleotide
XX CC sequences encoding them. A total of 40 PMM polypeptides (designated PMM
XX CC -1 to PMM-40) are disclosed. The sequences of the invention are useful
XX CC for diagnosing a condition or disease associated with the expression of
XX CC PMM in a subject, preparing a polyclonal or monoclonal antibody, and
XX CC generating an expression profile of a sample containing the
XX CC polynucleotides. The diseases or conditions associated with decreased
XX CC expression or overexpression of PMM are cell proliferation disorders
XX CC (e.g. cancer, atherosclerosis), neurological disorders (e.g. epilepsy,
XX CC Huntington's disease, stroke), immune/inflammatory disorders, (e.g. AIDS,
XX CC allergies), developmental disorders (e.g. hypothyroidism, Cushing's
XX CC syndrome), gastrointestinal or epithelial disorders, and infections. The
XX CC PMM polypeptides or their fragments are useful in screening compounds
XX CC for effectiveness as agonists or antagonists of the polypeptides, or in
XX CC altering the expression of the target polynucleotide and compounds that
XX CC specifically bind to, or modulate the activity of the polypeptide.
XX CC ABU92021-ABU92060 represent the human PMM polypeptides of the invention
XX SQ Sequence 824 AA;
Alignment Scores:
Pred. No.: 0 Length: 824
Score: 700.00 Matches: 700
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 68.0% Indels: 0
DB: Gaps: 0
US-10-825-632-2 (1-3120) x ABU92030 (1-824)
QY 760 CAACAACCTTTAAGGCCCAATCTAGTGGAAACTAGTGTCTCCCAACATACCGATGATCCA 819
Db 125 GlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMetAspPro 144
QY 820 AAATATATGCCCCGCTGATCCAGACTCGATGCTGCTTTTATACATAGCAACGATTTGGATA 879
Db 145 LysLeuCysProAlaAspProAspTyrIleAlaPheIleHisSerAsnAspIleTrpIle 164
QY 880 TCTAATCATCGTAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 939
Db 165 SerAsnIleValThrArgGluArgGluArgGluArgGluArgGluArgGluArgGluArg 184
```

QY 940 ATGGAAGAAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAAAGAAATTTGAT 999
Db 185 MetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluPheAsp 204
QY 1000 AGATATTCTGGCTATTGGTGGTGTCCAAAGAGCTGAAACAACTCCAGAGTGGTAAATTT 1059
Db 205 ArgTyrSerGlyTyrTrpCysProLysAlaGluThrThrProSerGlyGlyVal 224
QY 1060 CTTAGAAATTTCTATAGAAGAAATGATGAATCTGAGGTGGAAATTTATTCATGTTACATCC 1119
Db 225 LeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisValThrSer 244
QY 1120 CCTATGTGTGAACAAAGAGGCGAGATTCATTCCGTTATCTCTAAACAGGTACAGCAAT 1179
Db 245 ProMetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThrAlaAsn 264
QY 1180 CCTAAAGTCACTTTTAAGATGTCAGAAATATGATTGATGCTGAAGGAAGATCATAGAT 1239
Db 265 ProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIleAsp 284
QY 1240 GTCATAGATTAAGAACTAATTCACCTTTTGAGATTCATTGGAAGGAGTTGAATATTT 1299
Db 285 ValIleaspLysGluLeuLeuGlnProPheGluIleLeuPheGluGlyValGluTyrIle 304
QY 1300 GCCAGAGCTGGATCGACTCTCAGGAGAAATATGCTTGGTCCATCTCTACTAGATCGCTCC 1359
Db 305 AlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAspArgSer 324
QY 1360 CAGACTCGCTACAGATGTGTGATCTCACTCAATTAATTTATCCAGTAGAAGATGAT 1419
Db 325 GlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGluAspAsp 344
QY 1420 GTTATGGAAGCAGAGACTCAATTGAGTCAGTGCCTGATTTCTGTGACGCCACTAATTATC 1479
Db 345 ValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeuIleIle 364
QY 1480 TTATGAAGAAACACAGACATCTGGATAAATATCCATGACATCTTTCATGTTTTCCCAA 1539
Db 365 TyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPheProGln 384
QY 1540 AGTCACGAGAGAAATTTAGTTTATTTTGGCTCTGMAATGCAACACAGTTCCGTCAT 1599
Db 385 SerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPheArgHis 404
QY 1600 TTATCAAAATTTACATCTTATTTAAAGGAAGCAAAATATAACGATCCAGTGGTGGCTG 1659
Db 405 LeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGlyGlyLeu 424
QY 1660 CCTGCTCCAAGTATTTCAAGTGTCTATCAAGAGGATAGCAATTCACCGTGGTGAA 1719
Db 425 ProAlaProSerAspPheLysCysProIleLysGluIleAlaIleThrSerGlyGlu 444
QY 1720 TGGGAAGTTCTTGGCCGSCATGATCTAATATCCAAGTTGATGAAGTCAGAAGCTGCTGA 1779
Db 445 TrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArgLeuVal 464
QY 1780 TATTTGAAGGCACCAAGACTCCCTTTAGAGCATCACTGTGTAGTACGTTACGTA 1839
Db 465 TyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSerTyrVal 484
QY 1840 AATCTCGAGAGTGACAGGCTGACTGACCGTGGCTACTCACATTTCTTGTGTCATCAGT 1899
Db 485 AsnProGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCysIleSer 504
QY 1900 CAGCACTGTGACTCTTTATAGTATAGTATACCAAGAAATCCACACTGTGTGCTCC 1959
Db 505 GlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSer 524
QY 1960 CTTTACAGCTATCAAGTCTCAAGATGACCCAACTTGCACAAACAAAGGAATTTGGGCC 2019
Db 525 LeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPheTrpAla 544
QY 2020 ACCATTTTGGATTCAGCAGGTCCTCTCTCTGACTATACTCTCCAGAAATTTTCTCTTTT 2079

Db 545 ThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIlePheSerPhe 564
QY 2080 GAAAGTACTACTGGATTTACATTTGATGGATGCTCTACAGCCCTCATGATCTACAGCT 2139
Db 565 GluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnPro 584
QY 2140 GGAAGAAATATCCTACTGTGCTGTTTCATATATGTGTGCTCTCAGGTGCAGTTGGTCAAT 2199
Db 585 GlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeuValAsn 604
QY 2200 AATCGGTTTAAAGAGCTCAAGTATTTCCGCTTGATACCTAGCTCTCTAGGTATGTG 2259
Db 605 AsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrVal 624
QY 2260 GTTGTAGTGTAGACAAACAGGGATCCTGTCCACGAGGGCTTTAAATTTGAAGGCGCTTT 2319
Db 625 ValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPhe 644
QY 2320 AAATATAAATGGGTCAATAGAAATTCAGCATCAGGTGGAAGACTCCCAATATCTAGCT 2379
Db 645 LysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyrLeuAla 664
QY 2380 TCTCGATATGATTTCAATTCAGCTTAGATCGTGTGGGCATCCACGGCTGGTCTATGAGGA 2439
Db 665 SerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyrGlyGly 684
QY 2440 TACCTCTCCCTGATCGCATTTAATGCAGAGTCAAGATCTTCAGGGTTGCTATTGCTGGG 2499
Db 685 TyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAlaGly 704
QY 2500 GCCCAGTCACCTCTGTGATCTTCTATGATACAGGATACACGGAACGTTATATGGTGCAC 2559
Db 705 AlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMetGlyHis 724
QY 2560 CTGTACCAAGAAATCAACAGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAGTTC 2619
Db 725 ProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGluLysPhe 744
QY 2620 CCTCTGAACAAATCGTTTACTGCTCTTACATGGTTTCTGATGAGAAATGTCATTTT 2679
Db 745 ProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnValHisPhe 764
QY 2680 GCACATACCACTATATTACTGAGTTTTTTAGTCAGGGCTGGAAGCCATATCATTTACAG 2739
Db 765 AlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAspLeuGln 784
QY 2740 ATCTATCCTCAGGAGACACACAGATAAGAGTTTCCTGAATCGGAGAACATTATCAACTG 2799
Db 785 IleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyrGluLeu 804
QY 2800 CATCTTTTGCACTACCTTCAAGAAACCTTTGGATCACGTTATGCTCTCAAAAGTGA 2859
Db 805 HisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeuLysValIle 824
RESULT 11
ABB97361
ID ABB97361 standard; protein; 782 AA.
XX ABB97361;
AC
XX 27-JUN-2002 (first entry)
DT
XX
DE Novel human protein SEQ ID NO: 629.
XX
KW Human; antianaemic; vulnerary; antiinflammatory; immunomodulator;
KW antiinfertility; cerebroprotective; cytostatic; rheumatic; gene therapy;
KW neuroprotective; antiparkinsonian; protein therapy; EST;
XX expressed sequence tag.
OS Homo sapiens.
XX
XX WO200222660-A2.

XX 21-MAR-2003.
PD
XX
PF 10-SEP-2001; 2001WO-US026015.
XX
PR 11-SEP-2000; 2000US-00659671.
XX
XX (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
PI Xue AJ, Yang Y, Wehrman T, Drmanac RT;
XX
XX WPI; 2002-292408/33.
DR N-PSDB; ABN32547.
XX
PT An isolated polynucleotide for treating diseases associated with its
PT encoded polypeptide such as cancer and multiple sclerosis.
XX
PS Example 2; SEQ ID NO 629; 509pp; English.
XX
CC The present invention provides the protein and coding sequences of 444
CC novel human proteins. These were isolated from expressed sequences tags
CC (ESTs). They can be used to stimulate cell growth, to regulate
CC haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth
CC e.g. in burn treatment, to regulate the immune system e.g. to treat
CC multiple sclerosis, to regulate activin or inhibin e.g. to treat
CC infertility, to regulate haemostasis or thrombolysis e.g. to treat stroke
CC and cancer, to screen for drugs, to treat inflammatory conditions e.g.
CC rheumatoid arthritis, and to treat nervous system disorders e.g.
CC Parkinson's disease. The present sequence is a protein of the invention
XX
SQ Sequence 782 AA;

Alignment Scores:
Pred. No.: 0 Length: 782
Score: 658.00 Matches: 658
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 63.9% Indels: 0
DB: 5 Gaps: 0

US-10-825-632-2 (1-3120) x ABB97361 (1-782)

QY 214 ATGCAGCAGCAATGGAACAGAACAGCTGGGTGGTGGATTTGGAATTCGGGACTGT 273
Db 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20

QY 274 GAGCAGATATTGATCAGAGATCGGCTAAATTGGAGCCCTTTTATGTTGAGCGGTAT 333
Db 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40

QY 334 TCCTGGAGTCACTTAAAGCTTGTCCGATACCAGAAATATCATGGCTACATGATG 393
Db 41 SerTrpSerGlnLeuLysLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60

QY 394 GCTAAGGCACCATGATTTTCATGTTGTGAAGAGATGATCCAGATGGACCTCATTC 453
Db 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80

QY 454 GACAGAATCTATTACCTTGCATGCTGCTGGTGAAACAGAGAAATATACACTGTTTATCT 513
Db 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100

QY 514 GAAATTCCTCAAAATATCAATAGACGAGCTTAAATGCTCTCTTGGAGCCTCTTTTG 573
Db 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120

QY 574 GATCTTTTTCAGGCAACCTGGACTATGGAATGATTTCTCGAGAGAGAAGAACTATTAGA 633
Db 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuArg 140

QY 634 GAAAGAAACCATTTGAAACAGTCGGAATTCCTTACGATTATCCACCAAGGAAGTGA 693
Db

Db 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
QY 694 ACATTTCTGTTTCAAGCCGTAGTGAATTTATCATCGTAAAGATGGAGGCCCAAGGA 753
Db 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
QY 754 TTTACGCAACAACTTTAAGGCCCAATCTAGTGGAAACTAGTGTGTCCTCCAAACATACGATG 813
Db 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
QY 814 GATCCAAAATTTATCCCCGCTGATCCAGACTGGATTTGCTTTTATACATAGCAACGATAT 873
Db 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
QY 874 TGGATATCTAATCGTAACTACAGAGAAAGAGAGACTCATTTATGTGCAACATGAGCTA 933
Db 221 TrpIleSerAsnIleValThrArgGluGluArgGLeuThrTyrValHisAsnGluLeu 240
QY 934 GCCAATATGGAAGAGATGCCAGATCAGCTGAGTGGCTACCTTGTCTCCAGAGAA 993
Db 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu 260
QY 994 TTTGATAGATATCTGCTGCTATTGCTGCTGCCAAAGCTGAAACAACTCCACAGTGGTGT 1053
Db 261 PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly 280
QY 1054 AAAATTTCTAGAAATTTATATGAAGAAATGATGTAATCTGAGGTGGAAATTTATTCATGTT 1113
Db 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleHisVal 300
QY 1114 ACATCCCTATGTTGGAACCAAGGAGGCGAGATTCATTCGCTTATCTCTAAACAGGTACA 1173
Db 301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr 320
QY 1174 GCAAAATCCTAAAGTCACTTTTAAGATGTGAGAAATAATGATGCTGTAAGGAAGGATC 1233
Db 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
QY 1234 ATAGATGTCATAGTAAGGAACCTAAATCAACCTTTTGGATTTCTATTGAGGAGTTCAA 1293
Db 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
QY 1294 TATATTCAGAGCTGGATGCTCTGAGGGAATAATATGCTGTGCTCATCTACTAGAT 1353
Db 361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
QY 1354 CGCTCCAGAGCTCGCTTACAGATAGTGTGATCTCACCTGGAATTTATTTATCCAGTAGAA 1413
Db 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
QY 1414 GATGATGTTATGGAAGGCGAGAGACTCATGAGTCAGTCCCTGATTTCTGTGACGCCACTA 1473
Db 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
QY 1474 ATTATCTATGAAGAAACACAGACATCTGATAAATATCCATGACATCTTTCATGTTTTT 1533
Db 421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
QY 1534 CCCAAAGTCAAGAGGAAATGATTTATTTTGGCTCTGATGATGCAAAACAGGTTTC 1593
Db 441 ProGlnSerHisGluGluGluIlePheAlaSerGluCysLysThrGlyPhe 460
QY 1594 CGTCATTTATCAAAATATCATCTATTTTAAAGGAAGCAATATAACGATCCAGTGGT 1653
Db 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
QY 1654 GGGCTGCTGCTCCCAAGTGTATTTCAAGTGTCTTATTTTGGCTCTGATGATGCAAAACAGGTTTC 1713
Db 481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer 500
QY 1714 GGTGAATGGGAAGTCTTTGGCCGCGCATGATCTAATATCCAAAGTTGATGAAGTCAGAGG 1773
Db 501 GlyGluTrpGluValLeuGluArgHisGlySerAsnIleGlnValAspGluValArgArg 520

141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
QY 694 ACATTTCTGTTTCAAGCCGTAGTGAATTTATCATCGTAAAGATGGAGGCCCAAGGA 753
Db 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
QY 754 TTTACGCAACAACTTTAAGGCCCAATCTAGTGGAAACTAGTGTGTCCTCCAAACATACGATG 813
Db 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
QY 814 GATCCAAAATTTATCCCCGCTGATCCAGACTGGATTTGCTTTTATACATAGCAACGATAT 873
Db 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
QY 874 TGGATATCTAATCGTAACTACAGAGAAAGAGAGACTCATTTATGTGCAACATGAGCTA 933
Db 221 TrpIleSerAsnIleValThrArgGluGluArgGLeuThrTyrValHisAsnGluLeu 240
QY 934 GCCAATATGGAAGAGATGCCAGATCAGCTGAGTGGCTACCTTGTCTCCAGAGAA 993
Db 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu 260
QY 994 TTTGATAGATATCTGCTGCTATTGCTGCTGCCAAAGCTGAAACAACTCCACAGTGGTGT 1053
Db 261 PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly 280
QY 1054 AAAATTTCTAGAAATTTATATGAAGAAATGATGTAATCTGAGGTGGAAATTTATTCATGTT 1113
Db 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleHisVal 300
QY 1114 ACATCCCTATGTTGGAACCAAGGAGGCGAGATTCATTCGCTTATCTCTAAACAGGTACA 1173
Db 301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr 320
QY 1174 GCAAAATCCTAAAGTCACTTTTAAGATGTGAGAAATAATGATGCTGTAAGGAAGGATC 1233
Db 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
QY 1234 ATAGATGTCATAGTAAGGAACCTAAATCAACCTTTTGGATTTCTATTGAGGAGTTCAA 1293
Db 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
QY 1294 TATATTCAGAGCTGGATGCTCTGAGGGAATAATATGCTGTGCTCATCTACTAGAT 1353
Db 361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
QY 1354 CGCTCCAGAGCTCGCTTACAGATAGTGTGATCTCACCTGGAATTTATTTATCCAGTAGAA 1413
Db 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
QY 1414 GATGATGTTATGGAAGGCGAGAGACTCATGAGTCAGTCCCTGATTTCTGTGACGCCACTA 1473
Db 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
QY 1474 ATTATCTATGAAGAAACACAGACATCTGATAAATATCCATGACATCTTTCATGTTTTT 1533
Db 421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
QY 1534 CCCAAAGTCAAGAGGAAATGATTTATTTTGGCTCTGATGATGCAAAACAGGTTTC 1593
Db 441 ProGlnSerHisGluGluGluIlePheAlaSerGluCysLysThrGlyPhe 460
QY 1594 CGTCATTTATCAAAATATCATCTATTTTAAAGGAAGCAATATAACGATCCAGTGGT 1653
Db 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
QY 1654 GGGCTGCTGCTCCCAAGTGTATTTCAAGTGTCTTATTTTGGCTCTGATGATGCAAAACAGGTTTC 1713
Db 481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer 500
QY 1714 GGTGAATGGGAAGTCTTTGGCCGCGCATGATCTAATATCCAAAGTTGATGAAGTCAGAGG 1773
Db 501 GlyGluTrpGluValLeuGluArgHisGlySerAsnIleGlnValAspGluValArgArg 520

QY	1774	CTGGTATATTTTGAAGCACCACCAAGACTCCCTTTAGACATCACTGTAGTACGT	1833
Db	521	LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSer	540
QY	1834	TACGTAAATCTGGAGAGGTGACAGGCTGACTGACCGTGGCTACTCACATTTCTGTCTG	1893
Db	541	TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys	560
QY	1894	ATCAGTCAGCACTGTGACTCTTTTATAGTAAGTAACTAGTAAACCAAGAAATCCACACTGT	1953
Db	561	IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys	580
QY	1954	GTGTCCTTTTACAAGCTATCAAGTCTGAAGATGACCCCACTTGCAAAACAAAGGAATTT	2013
Db	581	ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe	600
QY	2014	TGGGCCACCATTTTGGATTCAGCAGGTCTCTTCTGACTATACCTCCAGAAATTTTC	2073
Db	601	TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe	620
QY	2074	TCTTTTGAAGTACTACTGGATTTACATTTGATGGATGCTCTCAAGCCCTCATGATCTA	2133
Db	621	SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu	640
QY	2134	CAGCTCGAAGAAATATCTACTGTGCTGTTTCATATATGCTGCTCAGGTG	2187
Db	641	GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnVal	658
RESULT 12			
ID	ABG61600	standard; protein; 658 AA.	
XX	ABG61600;		
AC	ABG61600;		
DT	12-AUG-2002	(first entry)	
DE	Human DPRP-1	splice variant #7.	
KW	Human; serine protease; dipeptidyl peptidase IV-related protein; DPRP;		
KW	DPPIV; infection; human immunodeficiency virus; HIV-1; HIV-2; pain;		
KW	diabetes; infertility; obesity; anorexia; Parkinson's disease; stroke;		
KW	heart failure; hypertension; urinary retention; osteoporosis; cancer;		
KW	ulcer; allergy; cancer; psychotic disorder; neurological disorder;		
KW	dykinesia; reproductive disorder; inflammatory disorder;		
KW	metabolic disorder.		
OS	Homo sapiens.		
XX	WO200231134-A2.		
XX	19-APR-2002.		
XX	12-OCT-2001; 2001WO-US031874.		
XX	12-OCT-2000; 2000US-0240117P.		
XX	(FERR) FERRING BV.		
XX	Qi S, Akinsanya KO, Riviere PJ, Junien J;		
XX	WPI; 2002-444178/47.		
XX	N-PSDB; ABK83331.		
XX	New dipeptidyl peptidase IV-related proteins and nucleic acids encoding		
XX	the proteins, useful for treating e.g. fungal, bacterial, protozoan and		
XX	viral infections, cancers, allergies, neurological disorders, or pain.		
XX	Disclosure; Page 70-72; 113pp; English.		
XX	The present invention relates to the isolation of novel human serine		
XX	proteases referred to as dipeptidyl peptidase IV (DPPIV)-related proteins		
XX	(DPRP). The dipeptidyl peptidase IV-related proteins (DPRP) and nucleic		
CC			

Db 241 AlaAsnMetGluGluuAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu 260
QY 994 TTTGATAGATATTCTGGCTATTGGTGGTGTCCAAAGCTGAACAACTCCAGTGGTGT 1053
Db 261 PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly 280
QY 1054 AATATCTTAGAATTTCTATATGCAAGAAATGATGAACTCAGCTGGAATTTATCATGTT 1113
Db 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal 300
QY 1114 ACATCCCTCATGTTGCAACAAAGAGGAGGACGATTCATTCCTGTTATCTTAAACAGGTACA 1173
Db 301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr 320
QY 1174 GCAATCTCTAAAGTCACCTTTTAAGATGTCAGAAATATGATGATGCTGAGGAGGATC 1233
Db 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
QY 1234 ATGATGTCATAGATAGGAACTAATTCACCTTTTGAGATTCCTATTGAAGGAGTTGAA 1293
Db 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluIleValGlu 360
QY 1294 TATATTGCCAGAGCTGGATGACTCTCTGAGGGAATAATGCTTGGTCCATCTCTAGAT 1353
Db 361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
QY 1354 CGTCCCGAGCTGCGCTACAGATAGTGTGATCTCAGCTGATTCCTGAGCCCACTA 1413
Db 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
QY 1414 GATGATGTTATGAAAGGAGAGAGCTCATGTGAGTCCGCTGATTCCTGAGCCCACTA 1473
Db 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
QY 1474 ATTATCTATCAAGAAACACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTTT 1533
Db 421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
QY 1534 CCCCAGAGTCACAAATATACATCTATTTTAAAGGAAAGCAATATAACGATCCAGTGGT 1593
Db 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
QY 1594 CGTCATTTATCAAAATATACATCTATTTTAAAGGAAAGCAATATAACGATCCAGTGGT 1653
Db 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
QY 1654 GGGCTGCTGCTCCAAAGTGATTTCAAGTCTCTATCAAGAGGAGATAGCAATACAGT 1713
Db 481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer 500
QY 1714 GGTGAATGGGAAGTCTTGCCCGCATGGATCTAATATCCAGTTGATGAAGTCAGAGG 1773
Db 501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520
QY 1774 CTGCTATATTTTGAAGCACCAGAGCTCCCTTTAGAGCATCACTGTAGTACGT 1833
Db 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSer 540
QY 1834 TACGTAAATCTCGAGAGGTGACAAAGCTGACATGACCGTGGCTACTCATCTTCTGCTGC 1893
Db 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
QY 1894 ATCAGTCAGACTGCTGCTCTTTATAGTAACTAGTAGTAAACAGAGAAATCCACACTGT 1953
Db 561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
QY 1954 GTGTCCTTTTCAAGCTATCAAGTCTGAAAGATGACCCCACTTTGCAAAACAAAGAAATTT 2013
Db 581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
QY 2014 TGGGCCACCAATTTTGGATTACAGAGGTCTCTTCTGACTATATCTCTCCAGAAATTTTC 2073

Db 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIlePhe 620
QY 2074 TCTTTTGAAGTACTACTTGGATTTTACATTTGATGGATGCTCTACAGCTCATATCTA 2133
Db 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
QY 2134 CAGCTCGGAAGAATATCTCTACTGCTGCTCTTCTATATATGTTGGT 2178
Db 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGly 655
RESULT 13
ABG61596
ID ABG61596 standard; protein; 661 AA.
XX AC ABG61596;
XX XX 12-AUG-2002 (first entry)
XX DE Human DPPP-1 splice variant #3.
XX XX Human; serine protease; dipeptidyl peptidase IV-related protein; DPPP;
KW DPPIV; infection; human immunodeficiency virus; HIV-1; HIV-2; pain;
KW diabetes; infertility; obesity; anorexia; Parkinson's disease; stroke;
KW heart failure; hypertension; urinary retention; osteoporosis; cancer;
KW ulcer; allergy; cancer; psychotic disorder; neurological disorder;
KW dyskinesia; reproductive disorder; inflammatory disorder;
KW metabolic disorder.
XX OS Homo sapiens.
XX XX WO200231134-A2.
XX PD 18-APR-2002.
XX PF 12-OCT-2001; 2001WO-US031874.
XX PR 12-OCT-2000; 2000US-0240117P.
XX PA (FERR) FERRING BV.
XX PI Qi S, Akinsanya KO, Riviere PJ, Junien J;
XX WPI; 2002-444178/47.
DR N-PSDB; ABK83327.
XX New dipeptidyl peptidase IV-related proteins and nucleic acids encoding
PT the proteins, useful for treating e.g. fungal, bacterial, protozoan and
PT viral infections, cancers, allergies, neurological disorders, or pain.
XX PS Disclosure; Page 63-65; 113pp; English.
XX The present invention relates to the isolation of novel human serine
CC proteases referred to as dipeptidyl peptidase IV (DPPIV)-related proteins
CC (DPPP). The dipeptidyl peptidase IV-related proteins (DPPP) and nucleic
CC acids encoding them are useful for treating infections such as fungal,
CC bacterial, protozoan and viral infections, particularly infections caused
CC by human immunodeficiency virus (HIV-1 or HIV-2), pain, diabetes,
CC precocious puberty, infertility, obesity, anorexia, bulimia, Parkinson's
CC disease, acute heart failure, hypotension, hypertension, urinary
CC retention, osteoporosis, angina pectoris, stroke, ulcers, asthma,
CC allergies, cancers, migraine, vomiting, psychotic and neurological
CC disorders (e.g. anxiety, dementia, or schizophrenia), and dyskinesias.
CC These may also be used in discovering therapeutic agents for the
CC treatment of reproductive, inflammatory and metabolic disorders. ABG61591
CC -ABG61612 represent human DPPP proteins
XX Sequence 661"AA;
SQ Alignment Scores:
Pred. No.: 0 Length: 661
Score: 655.00 Matches: 655
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0

Query Match:	63.6%	Indels:	0
DB:	5	Gaps:	0
US-10-825-632-2 (1-3120) x ABG61596 (1-661)			
QY	214	ATGGCAGCAGCAATATGGAACACAGACAGCTGGGTGTTGAGATATTGAAACTGCGGACTGT	273
DB	1	MetAlaAlaMetGluThrGluGlnLeuGluValGluLeuPheGluThrAlaAspCys	20
QY	274	GAGAGAAATATTGAATCACAGGATCGGCTAAATTTGAGCCTTTTATGTTGAGCGGTAT	333
DB	21	GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheYrValGluArgYr	40
QY	334	TCCTGGAGTCAGCTTAAAGCTGCTGCGATACAGAAATATCATGCTCATCATG	393
DB	41	SerTrpSerGlnLeuLysLysLeuAlaAspThrArgLysYrHisGlyYrMetMet	60
QY	394	GCTAAGCACCACATGATTTTCATGTTTGTGAAGAGAAATGATCCAGATGGACCTCATTTCA	453
DB	61	AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer	80
QY	454	GACGAATCTATTACCTTGCCTGCTGTGTGAGAACAGAGAAATACACTGTTTATTC	513
DB	81	AspArgIleYrYrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheYrSer	100
QY	514	GAATTTCCCAAACTATCAATAGACAGCAGTCTTAATGCTCTCTTGGAAAGCCTCTTTTG	573
DB	101	GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu	120
QY	574	GATCTTTTTCAGCAACACTGGACTATGGAATGATTTCTCGAGAAAGAACTATTAAAG	633
DB	121	AspLeuPheGlnAlaThrLeuAspYrGlyMetYrSerArgGluGluLeuLeuLeuArg	140
QY	634	GAAGAAGCGCATTTGGAACAGTCGGAATTTGCTTTACGATATTATCAACAAAGAACTGGA	693
DB	141	GluArgLysArgIleGlyThrValGlyIleAlaSerYrAspYrHisGlnGlySerGly	160
QY	694	ACATTTCTGTTTCAAGCCCGTAGTGGAAATTTATCACGTAAAGATGAGGGCCACAGCA	753
DB	161	ThrPheLeuPheGlnAlaGlySerGlyIleYrHisValLysAspGlyGlyProGlnGly	180
QY	754	TTTACCCCAACACCTTTAAGGCCCAATCTAGTGAACACTAGTGTCCACATACCGATG	813
DB	181	PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet	200
QY	814	GATCCAAATTTATGCCCGCTGATCCAGACTGGATTCCTTTTATACATAGCAACGATATT	873
DB	201	AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle	220
QY	874	TGGATATCTAATCGTAAACAGAGAGAAAGGAGACTCACTTATGTGCACAAATGAGCTA	933
DB	221	TrpIleSerAsnIleValThrArgGluGluArgLeuThrYrValHisAsnGluLeu	240
QY	934	GCCACATGGAAGAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAAGAGAA	993
DB	241	AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu	260
QY	994	TTTGATAGATATCTGGCTATTGTTGGTGTCCAAAAGCTGAAACAACTCCCAAGTGGTGT	1053
DB	261	PheAspArgYrSerGlyYrTrpTrpCysProLysAlaGluThrThrProSerGlyGly	280
QY	1054	AAAATCTTGAATTTCTATGAAAGAAATGATGAATCTGAGTGGAAATATTTCATGTT	1113
DB	281	LysIleLeuArgIleLeuYrGluGluAsnAspGluSerGluValGluIleIleHisVal	300
QY	1114	ACATCCCTATGTTGGAAACAGGAGGCGAGATTTCATTCGTTTATCTTAAACACAGTACA	1173
DB	301	ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgYrProLysThrGlyThr	320
QY	1174	GCAAATCCCTAAAGTCACCTTTTAAGATGTCAGAAATATGATGATGCTGAGGAGGATC	1233
DB	321	AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle	340

QY	1234	ATAGATGTCATAGATAAGAACTAATTCACCTTTTGGAGATCTATTATTTGAAGAGTTGAA	1293
DB	341	IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu	360
QY	1294	TATATTGCCAGAGCTGGATGAGACTCTCTGAGGGAATAATATGCTTGGCTCCTACTAGAT	1353
DB	361	TyrIleAlaArgAlaGlyTrpThrProGluGlyLysYrAlaTrpSerIleLeuLeuAsp	380
QY	1354	CGCTCCAGACTCGCTTACAGATAGTGTGTGATCTCACCTGAAATATTATTTATCCCAAGTAGAA	1413
DB	381	ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu	400
QY	1414	GATCATGTTATGAAAGGAGAGACTCATTGAGTCAGTGCCTGATTCCTGTCACCCACTA	1473
DB	401	AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu	420
QY	1474	ATTATCTATCAAGAAACAAACAGACATCTGGATAAATATCCATGACATCTTCATGTTT	1533
DB	421	IleIleYrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe	440
QY	1534	CCCCAAAGTCACGAAGAGAAATTCAGTTTATTTTGCCTCTCTGAATCGMAAACAGGTTTC	1593
DB	441	ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe	460
QY	1594	CGTCATTTATACAAAATTCATCTATTTTAAAGGAAGCAAAATATAACGATCCAGTGGT	1653
DB	461	ArgHisLeuYrLysIleThrSerIleLeuLysGluSerLysYrLysArgSerSerGly	480
QY	1654	GGCTCGCTGCTCCAGTGATTTCAAGTGTCTCTATCAAGAGGAGATAGCAATTTACCACT	1713
DB	481	GlyLeuProAlaProSerAspPheLysCysProIleLysGluIleAlaIleThrSer	500
QY	1714	GGTGAATGGGAATGTTCTTGGCCGCGCATGGATCTTAATATCCAAAGTTGATGAAGTCAGAAGG	1773
DB	501	GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArg	520
QY	1774	CTGGTATATTTTGAAGGCCAACCAAGACTCCCTTTTAGAGCATCACCTGTACGTAGTCACT	1833
DB	521	LeuValYrPheGluGlyThrLysAspSerProLeuGluHisHisLeuYrValValSer	540
QY	1834	TACGTAATTCCTGGAGAGGTGACCAAGCTGACTCACCGTGGCTACTCACATCTTCTGCTGC	1893
DB	541	TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyYrSerHisSerCysCys	560
QY	1894	ATCAGTCAGCACTGTGACTTCTTTTATAGTAAGTATAGTAACCAAGAGAAATCCACACTCT	1953
DB	561	IleSerGlnHisCysAspPheIleSerLysYrSerAsnGlnLysAsnProHisCys	580
QY	1954	GTGTCCCTTTTACAACTATCAAGTCCCTGAAGATGACCCAACTTGCACAAACAGGAATTT	2013
DB	581	ValSerLeuYrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe	600
QY	2014	TGGGCCACCATTTTGGATTGAGAGGCTCTTCTTCTGACTATATCTCTCCAGAAATTTTC	2073
DB	601	TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspYrThrProGluIlePhe	620
QY	2074	TCCTTTTGAAGTACTGCTGGAATTTACATTTGATGGATGCTCTACAGCTCATGATCTA	2133
DB	621	SerPheGluSerThrThrGlyPheThrLeuYrGlyMetLeuYrLysProHisAspLeu	640
QY	2134	CAGCTCGAAGAAATATCTACTGCTGCTTTCATATATGCTGCT 2178	
DB	641	GlnProGlyLysLysYrProThrValLeuPheIleYrGlyGly 655	

RESULT 14

ABG61594

ID ABG61594 standard; protein; 690 AA.

XX ABG61594;

XX

XX 12-AUG-2002 (first entry)

XX Human DPRP-1 splice variant #1.

DE

XX Human; serine protease; dipeptidyl peptidase IV-related protein; DPPP;
KW DPPIV; infection; human immunodeficiency virus; HIV-1; HIV-2; pain;
KW Diabetes; infertility; obesity; anorexia; Parkinson's disease; stroke;
KW heart failure; hypertension; urinary retention; osteoporosis; cancer;
KW ulcer; allergy; cancer; psychotic disorder; neurological disorder;
KW dyskinesia; reproductive disorder; inflammatory disorder;
KW metabolic disorder.
XX Homo sapiens.
XX W0200231134-A2.
XX 18-APR-2002.
XX 12-OCT-2001; 2001WO-US0311874.
XX 12-OCT-2000; 2000US-0240117P.
XX (FERR) FERRING BV.
XX Qi S, Akinsanya KO, Riviere PJ, Junien J;
XX WPI; 2002-444178/47.
XX N-PSDB; ABK83325.
XX New dipeptidyl peptidase IV-related proteins and nucleic acids encoding
PT the proteins, useful for treating e.g. fungal, bacterial, protozoan and
PT viral infections, cancers, allergies, neurological disorders, or pain.
XX Disclosure; Page 59-61; 113pp; English.
XX The present invention relates to the isolation of novel human serine
CC proteases referred to as dipeptidyl peptidase IV (DPPIV)-related proteins
CC (DPPP). The dipeptidyl peptidase IV-related proteins (DPPP) and nucleic
CC acids encoding them are useful for treating infections such as fungal,
CC bacterial, protozoan and viral infections, particularly infections caused
CC by human immunodeficiency virus (HIV-1 or HIV-2), pain, diabetes,
CC precocious puberty, infertility, obesity, anorexia, bulimia, Parkinson's
CC disease, acute heart failure, hypotension, hypertension, urinary
CC retention, osteoporosis, angina pectoris, stroke, ulcers, asthma,
CC allergies, cancers, migraine, vomiting, psychotic and neurological
CC disorders (e.g. anxiety, dementia, or schizophrenia), and dyskinesias.
CC These may also be used in discovering therapeutic agents for the
CC treatment of reproductive, inflammatory and metabolic disorders. ABG61591
CC -ABG61612 represent human DPPP proteins
XX
SQ Sequence 690 AA;

Alignment Scores:
Pred. No.: 0 Length: 690
Score: 655.00 Matches: 655
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 63.6% Indels: 0
DB: 5 Gaps: 0

US-10-825-632-2 (1-3120) x ABG61594 (1-690)
QY 214 ATGGCAGCAGCAATGGAAACAGACAGCTGGGTGGTGGATATTGAACTGCGGACTGT 273
Db 1 MetAlaAlaMetGluThrGluGluLeuGluValGluLeuPheGluThrAlaAspCys 20
QY 274 GNGGGAATATTGAATCACAGGATCGGCCTAAATTTGGAGCCTTTTATGTTGAGCGGTAT 333
Db 21 GluGluAenIIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
QY 334 TCCTGGAGTCAGCTTAAAGAGCTGCTTGGCCGATACAGAAAATATCATGGCTACATGATG 393
Db 41 SerTrpSerGlnLeuLysLysLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
QY 394 GCTAAGGACCAATGATTTTCATGTTGTGAAGGAATGATCCAGATGGACCTCATTC 453

Db 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
QY 454 GACAGAACTATTACCTTGCATCTCTGTGAGAAACAGAGAAATACACTGTTTATTCT 513
Db 81 AspArgileTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
QY 514 GAAATTTCCAAACATCATAGAGCAGCTCTTAATGCTCTCTTGGAAAGCCCTCTTTG 573
Db 101 GluileProLysThrIleAsnArgAlaValLeuMetLeuSerTrpLysProLeuLeu 120
QY 574 GATCTTTTCAGCAACACTGGACTGATGATGATGATGATGATGATGATGATGATGAT 633
Db 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg 140
QY 634 GAAAGAAAACGCAATTTAGGCCCAATCTAGTGGAACTAGTGGTCTCCCAACATACGATG 693
Db 141 GluArgLysArgileGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
QY 694 ACATTTCTGTTTCAAGCCGGTAGTGGAAATTTATCACGTAAAGATGGAGGCCACAGGA 753
Db 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
QY 754 TTTACGCAACAACTTTAAGGCCCAATCTAGTGGAACTAGTGGTCTCCCAACATACGATG 813
Db 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
QY 814 GATCCAAATTTATGCCCGCTGATCCAGACTGGATTTGCTTTTATACATAGCAACGATATT 873
Db 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
QY 874 TGGATATCTAAACATCGTAAACCAGAGAAGAGGAGACTCCTTATGTGCAATGAGCTA 933
Db 221 TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu 240
QY 934 GCCAATGAGAAAGATGCCAGATCCAGATCAGCTGGAGTCGCTACTTGTCTCCAGAGAGAA 993
Db 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu 260
QY 994 TTTGATGATATTCTCGCTATTGTTGGTGTCCAAAGCTGAAACAACTCCCAAGTGGTGGT 1053
Db 261 PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly 280
QY 1054 AAAATTTCTAGAAATTTCTATATGAAGAAATGATGAATCTGAGGTGGAAATTTATCATGTT 1113
Db 281 LysileLeuArgileLeuTyrGluGluAsnAspGluSerGluValGluIleHisVal 300
QY 1114 ACATCCCTATGTTGAAACAAAGAGGCGAGATTCATTCCTGTTCTTAAACAGAGTACA 1173
Db 301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr 320
QY 1174 GCMAATCCTAAAGTCACTTTTAAGATGTCAGAAATAATGATGATGCTGGAAGGAGATC 1233
Db 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
QY 1234 ATAGATGTCATAGATAAGAACTAAATTCACCTTTTGGAGATTCATTTTGAAGGAGTTGAA 1293
Db 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
QY 1294 TATATTCAGAGCTGAGTGGACTCCTGAGGGAATAATGCTGCTGCTCCTACTAGAT 1353
Db 361 TyrIleAlaArgAlaGlyTyrThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
QY 1354 CGCTCCAGACTCGCTTACAGATAGTGTGATCTACCTGAAATTTATTTATCCAGTAGAA 1413
Db 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
QY 1414 GATGATGTTATGGAAGAGCAGAGACTCATGTGAGTCCAGTCCCTGATTCCTGTGAGCGCATA 1473
Db 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
QY 1474 ATTATCTATGAAGAAACAAACAGACATCTGATTAATATCCATGACATCTTTCTGTTT 1533
Db 421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440

Db	161	ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly	180
Qy	754	TTTACGCAACACCTTTAAGCCCAATCTAGTGGAACTAGTTGTCCCAACATACGATG	813
Db	181	PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet	200
Qy	814	GATCCAAATATGCCCCGCTGATCCAGACTGGATTGCTTTTATACATAGCAACGATATT	873
Db	201	AspProLysLeuCysProAlaAspProAspTrrPileAlaPheIleHisSerAsnSpile	220
Qy	874	TGGATATCTTAACATCGTAACACAGAGAAAGAGACTCACTTATGTGCACAAATGAGCTA	933
Db	221	TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu	240
Qy	934	GCCAAACATGGAAGAAGATGCCAGATCAGCTGGAGTGGCTTGTCTCCAAAGAAGAA	993
Db	241	AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu	260
Qy	994	TTTGATAGATATTCTGGCTATTGGTGGTGTCCAAAGCTGAAACAACTCCAGTGGTGGT	1053
Db	261	PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly	280
Qy	1054	AAATTTCTTAGAATTTCTATATGAAAGAAATGATGAATCTGAGGTGGAAATTTATCATGTT	1113
Db	281	LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal	300
Qy	1114	ACATCCCTATGTTGGAACAAGAGGAGGAGATTCATCCGTTATCTTAAACAGGTACA	1173
Db	301	ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr	320
Qy	1174	GCAATCTCTAAAGTCACCTTTTAAGATGTGAGAAATAATGATTGATGCTGGAAGGAGATC	1233
Db	321	AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle	340
Qy	1234	ATAGATGTCATAGATAAGAACTAATTCACCTTTTGAGATTCTATTGGAAGGAGTTGAA	1293
Db	341	IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu	360
Qy	1294	TATATTGCCAGAGCTGGATGGACTCTCTGAGGGAATAATGCTTGGTCCATCCTACTAGAT	1353
Db	361	TyrIleAlaArgAlaGlyTyrThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp	380
Qy	1354	CGTCCCGAGCTCCCTACAGATAGTGTGATCTACCTGAATTTATTTCCAGTAGAA	1413
Db	381	ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu	400
Qy	1414	GATGATGTTTGAAGAGCAGAGACTCATTGAGTCAGTCCTGATTCGTGACGCCACTA	1473
Db	401	AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu	420
Qy	1474	ATTATCTATGAAGAAACAACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTTTT	1533
Db	421	IleIleIleTyrGluGluThrThrAspIleTrrPileAsnIleHisAspIlePheHisValPhe	440
Qy	1534	CCCCAAGTCACGAAGAGGAAATTTAGTTTATTTTTCCTCTGAATGCAAAACAGGTTTC	1593
Db	441	ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe	460
Qy	1594	CGTCATTTTATACAAATTTACATCTATTTTTAAAGGAAGCAAAATATAACGATCCAGTGGT	1653
Db	461	ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly	480
Qy	1654	GCGTCGCTGCTCCAGTGATTTCAAGTGTCCCTATCAAAGAGAGATAGCAATTTACCACT	1713
Db	481	GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer	500
Qy	1714	GGTGAATGGAGTTCTTGGCCGATGGATCTTAATATCCAAGTTGATGAAGTCAGAGG	1773
Db	501	GlyGluTrrPrrGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg	520
Qy	1774	CTGGTATATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACCCTGTACGTAGTCAGT	1833

Db	521	LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSer	540
Qy	1834	TACGTAAATCTCGAGAGGTGACAGGCTGACTGACCGTGGCTACTCACATTTCTTGTCTGC	1893
Db	541	TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys	560
Qy	1894	ATCAGTCAGCACTGTGCACTTCTTTTATAGTAAGTATAGTAACCAAGAAATCCACACTGT	1953
Db	561	IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys	580
Qy	1954	GTGTCCCTTTTACAAAGCTATCAAGTCTCTGAAGATGACCCCACTTGCAAAACAAAGAAATTT	2013
Db	581	ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe	600
Qy	2014	TGGGCCACCATTTTGGATTCA	2034
Db	601	TrpAlaThrIleLeuAspSer	607

Search completed: May 2, 2006, 03:26:17
Job time : 525 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: May 2, 2006, 03:26:34 ; Search time 20.8 Seconds
(without alignments)
2886.503 Million cell updates/sec

Title: US-10-825-632-2
Perfect score: 1030
Sequence: 1 aagtgtaagctccgagg.....agaattactaaaaaaaaa 3120

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 60.0 , Fgapext 7.0
Delop 60.0 , Delext 60.0

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 565918

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Listing first 45 summaries

Command line parameters:
-MODEL=frame+n2p.model -DEV=xlp
-O=/abss/ABSSWEB.spool/US10825632/runat_01052006_110237_4459/app_query.fasta_1
-DB=PIR -QFMT=fastan -SUFFIX=oligo.n2p.rpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc
-NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abs02p
-USER=US10825632 @CGN 1.1_63 @runat_01052006_110237_4459 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=60 -FGAPEXT=7
-YGAPOP=60 -YGAPEXT=60 -DELOP=60 -DELEXT=60

Database : PIR_90:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES				Description	
Result No.	Query	Score	Match Length DB ID		
1	9	0.9	249	2	T09139
C 2	9	0.9	297	2	H84168
C 3	9	0.9	667	2	S63587
C 4	9	0.9	721	2	F70649
C 5	9	0.9	2073	1	EWASBE
C 6	9	0.9	2194	1	GNVYE7
C 7	8	0.8	65	2	AH2228
C 8	8	0.8	85	1	FDFL4W
C 9	8	0.8	91	2	A22592
C 10	8	0.8	97	2	S02376
C 11	8	0.8	113	2	B75632
C 12	8	0.8	123	2	F64083
C 13	8	0.8	131	1	SEPG
C 14	8	0.8	141	2	F38164

15	8	0.8	145	2	AE2962
16	8	0.8	150	2	S09872
17	8	0.8	156	2	S77503
18	8	0.8	166	2	T49489
19	8	0.8	180	2	A98321
20	8	0.8	183	2	A72657
21	8	0.8	200	2	GB4822
22	8	0.8	203	2	JC6113
23	8	0.8	205	2	S70618
24	8	0.8	216	2	B26423
25	8	0.8	216	2	A61259
26	8	0.8	225	1	MMIH68
27	8	0.8	244	2	T00449
28	8	0.8	247	2	AG3513
29	8	0.8	250	2	C87443
30	8	0.8	264	2	E95410
31	8	0.8	275	2	S21348
32	8	0.8	301	2	H95308
33	8	0.8	303	2	D84348
34	8	0.8	308	2	C83586
35	8	0.8	310	2	A60967
36	8	0.8	317	2	I46916
37	8	0.8	321	2	T12497
38	8	0.8	330	1	H69798
39	8	0.8	342	2	H70653
40	8	0.8	346	2	T13973
41	8	0.8	346	2	T14033
42	8	0.8	346	2	T14049
43	8	0.8	346	2	T13835
44	8	0.8	346	2	T17105
45	8	0.8	346	2	T13834

ALIGNMENTS

RESULT 1
T09139
26S proteasome alpha chain - spinach
C:Species: Spinacia oleracea (spinach)
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
C:Accession: T09139
R:Rito, N.; Tomizawa, K.; Tanaka, K.; Matsui, M.; Kendrick, R.; Sato, T.; Nakagawa, H.
Plant Mol. Biol. 34, 307-316, 1997
A:Title: Characterization of 26S proteasome alfa- and beta-type and ATPase subunits from
A:Reference number: Z16562; MUID:97351564; PMID:9207846
A:Accession: T09139
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-249 <ITO>
A:Cross-references: UNIPROT:O24362; UNIPARC:UPI0000132450; EMBL:D78173; NID:G2285801; PII
C:Superfamily: multicatalytic endopeptidase complex chain C9

Alignment Scores:
Pred. No.: 6.39 Length: 249
Score: 9.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 0.9% Indels: 0
DB: 2 Gaps: 0

US-10-825-632-2 (1-3120) x T09139 (1-249)

QY 17 GAGGCCAAGCGCGTGTACTGCGCC 43
|||||
Db 234 GUAALYSALAAAlaAlaThrAlaAla 242

RESULT 2

H84168
molybdenum cofactor biosynthesis protein [imported] - Halobacterium sp. NRC-1
C:Species: Halobacterium sp. NRC-1
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C:Accession: H84168
R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.

Leithausen, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jabbid
Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li
A:Title: Genome sequence of Halobacterium species NRC-1.
A:Reference number: A84160; MUID:20504483; PMID:11016950
A:Accession: H84168
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-297 <STO>
A:Cross-references: UNIPROT:Q9HST6; UNIPARC:UPI0000063541; GB:AE004437; NID:G10579734; E
C:Genes: moaE
A:Gene: moaE

Alignment Scores:
Pred. No.: 6.22 Length: 297
Score: 9.00 Matches: 9
Percent Similarity: 100.0% Conservations: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 0.9% Indels: 0
DB: 2 Gaps: 0

US-10-825-632-2 (1-3120) x H84168 (1-297)

QY 48 GCAGCGCGCAGTAGCAGCGCCTTG 22
DB 50 AlaAlaAlaValAlaAlaAlaLeu 58

RESULT 3
S63587
gene pacc protein - Aspergillus niger
C:Species: Aspergillus niger
C:Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C:Accession: S63587
R:MacCabe, A.P.; van den Hombergh, J.P.T.W.; Tilburn, J.; Arst Jr., H.N.; Visser, J.
Mol. Gen. Genet. 250, 367-374, 1996
A:Title: Identification, cloning and analysis of the Aspergillus niger gene pacc, a wide
A:Reference number: S63587; MUID:96180854; PMID:8602152
A:Accession: S63587
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-667 <MAC>
A:Cross-references: UNIPROT:Q00203; UNIPARC:UPI000006C82D; GB:X98417; GB:S81349; NID:G14
C:Genetics:
A:Gene: pacc

Alignment Scores:
Pred. No.: 5.47 Length: 667
Score: 9.00 Matches: 9
Percent Similarity: 100.0% Conservations: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 0.9% Indels: 0
DB: 2 Gaps: 0

US-10-825-632-2 (1-3120) x S63587 (1-667)

QY 26 GCGCTGCTACTGCGCGCCTGCTTCT 52
DB 43 AlaAlaAlaThrAlaAlaAlaSer 51

RESULT 4
F70649
probable Acyl-CoA Dehydrogenase - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C:Accession: F70649
R:Cole, S.T.; Brooch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.
i: Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the Biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230

A:Accession: F70649
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-721 <COL>
A:Cross-references: UNIPROT:P95097; UNIPARC:UPI00000D600B; GB:Z83866; GB:AL123456; NID:G
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: fadE22

Alignment Scores:
Pred. No.: 5.41 Length: 721
Score: 9.00 Matches: 9
Percent Similarity: 100.0% Conservations: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 0.9% Indels: 0
DB: 2 Gaps: 0

US-10-825-632-2 (1-3120) x F70649 (1-721)

QY 45 GCGCGCGCAGTAGCAGCGCCTTGCC 19
DB 284 AlaAlaAlaValAlaAlaAlaLeuAla 292

RESULT 5
BWASBE
bimE protein - Emericella nidulans
C:Species: Emericella nidulans, Aspergillus nidulans
C:Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 09-Jul-2004
C:Accession: A37879
R:Engle, D.B.; Omani, S.A.; Omani, A.H.; Roeborough, S.; Xiang, X.; Morris, N.R.
J. Biol. Chem. 265, 16132-16137, 1990
A:Title: A negative regulator of mitosis in Aspergillus is a putative membrane-spanning I
A:Reference number: A37879; MUID:90375468; PMID:1697851
A:Accession: A37879
A:Molecule type: mRNA
A:Residues: 1-2073 <ENG>
A:Cross-references: UNIPROT:P24686; UNIPARC:UPI0000126941; GB:M59705; GB:J05607; NID:G16
A:Note: In addition to three predicted transmembrane domains, there are several potential
aspartic kinase, and one sequence that resembles a nuclear localization signal
C:Comment: This protein is part of a regulatory pathway that includes the nimA protein k
ter mitosis and prevent them from leaving mitosis.
C:Genetics:
A:Gene: bimE
C:Superfamily: bimE protein
C:Keywords: cell cycle control; mitosis; transmembrane protein
F:1623-1643/Domain: transmembrane #status predicted <TM1>
F:1685-1703/Domain: transmembrane #status predicted <TM2>
F:1746-1764/Domain: transmembrane #status predicted <TM3>

Alignment Scores:
Pred. No.: 4.58 Length: 2073
Score: 9.00 Matches: 9
Percent Similarity: 100.0% Conservations: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 0.9% Indels: 0
DB: 1 Gaps: 0

US-10-825-632-2 (1-3120) x BWASBE (1-2073)

QY 754 ATCCTTGCGCCCTCCATCTTTACGT 728
DB 960 IleLeuValAlaLeuHisLeuLeuArg 968

RESULT 6
GNVEY7
genome polyprotein - human enterovirus 70 (strain J670/71)
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core prot
NA-directed RNA polymerase (NC 2.7.7.48)
C:Species: human enterovirus 70
C:Date: 30-Jun-1993 #sequence_revision 30-Jun-1993 #text_change 09-Jul-2004
C:Accession: A36253
R:Ryan, M.D.; Jenkins, O.; Hughes, P.J.; Brown, A.; Knowles, N.J.; Booth, D.; Minor, P.D.
J. Gen. Virol. 71, 2291-2299, 1990

A:Title: The complete nucleotide sequence of enterovirus type 70: relationships with other
A:Reference number: A36253; MUID:91037960; PMID:2172447

A:Accession: A36253
A:Molecule type: genomic RNA
A:Residues: 1-2194 <RYA>
A:CROSS-references: UNIPROT:P32537; UNIPARC:UPI00001749F5; GB:D00820
C:Superfamily: poliovirus genome polyprotein
C:Keywords: coat protein; core protein; genome-linked protein; hydrolase; nucleotidyltransferase
F:1-69/Product: coat protein 1A #status predicted <VP4>
F:70-319/Product: coat protein 1B #status predicted <VP2>
F:320-561/Product: coat protein 1C #status predicted <VP3>
F:562-871/Product: coat protein 1D #status predicted <VP1>
F:872-1014/Product: coat protein 2A #status predicted <P2A>
F:1015-1113/Product: coat protein 2B #status predicted <P2B>
F:1114-1443/Product: coat protein 2C #status predicted <P2C>
F:1444-1532/Product: protein 3A #status predicted <P3A>
F:1533-1554/Product: genome-linked protein VPg #status predicted <VPG>
F:1555-1737/Product: proteinase 3C #status predicted <P3C>
F:1738-2194/Product: RNA-directed RNA polymerase #status predicted <RRP>
F:1535/Binding site: phosphoryl-RNA (Tyr) (covalent) #status predicted

Alignment Scores:
Pred. No.: 4.54 Length: 2194
Score: 9.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 0.9% Indels: 0
DB: 1 Gaps: 0

US-10-825-632-2 (1-3120) x GNNYE7 (1-2194)

QY 1022 GTCCAAAGCTGAACACTCCAGTG 1048

DB 1160 ValGlnLysLeuLysGlnLeuProVal 1168

RESULT 7

AH2228

hypothetical protein asl3383 [imported] - Nostoc sp. (strain PCC 7120)

C:Species: Nostoc sp. PCC 7120

A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004

C:Accession: AH2228

R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,

Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S.

DNA Res. 8, 205-213, 2001

A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena

A:Reference number: AB1807; MUID:21595285; PMID:11759840

A:Accession: AH2228

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-65 <KUR>

A:CROSS-references: UNIPROT:Q8VRR0; UNIPARC:UPI000000CE789; GB:BA0000019; PIDN:BA075082.1;

A:Experimental source: strain PCC 7120

C:Genetics:

A:Gene: asl3383

Alignment Scores:

Pred. No.: 85.7 Length: 65

Score: 8.00 Matches: 8

Percent Similarity: 100.0% Conservative: 0

Best Local Similarity: 100.0% Mismatches: 0

Query Match: 0.8% Indels: 0

DB: 2 Gaps: 0

US-10-825-632-2 (1-3120) x AH2228 (1-65)

QY 2866 AAAATTATACCTTTTAGACGAG 2843

DB 37 LysIleIleSerLeuLeuGluGln 44

RESULT 8

FDL4W

antifreeze protein 4 precursor - winter flounder

C:Species: Pseudopleuronectes americanus (winter flounder)

C:Date: 01-Sep-1981 #sequence_revision 01-Sep-1981 #text_change 09-Jul-2004

C:Accession: A03193

R:Lin, Y.; Gross, J.K.

Proc. Natl. Acad. Sci. U.S.A. 78, 2825-2829, 1981

A:Title: Molecular cloning and characterization of winter flounder antifreeze cDNA.

A:Reference number: A03193; MUID:81247379; PMID:6265915

A:Accession: A03193

A:Molecule type: mRNA

A:Residues: 1-85 <LIN>

A:CROSS-references: UNIPROT:P02734; UNIPARC:UPI0000174278

C:Superfamily: antifreeze protein

C:Keywords: antifreeze

F:1-21/Domain: signal sequence #status predicted <SIG>

F:22-85/Product: antifreeze protein 4 #status predicted <MAT>

Alignment Scores:

Pred. No.: 82.2 Length: 85

Score: 8.00 Matches: 8

Percent Similarity: 100.0% Conservative: 0

Best Local Similarity: 100.0% Mismatches: 0

Query Match: 0.8% Indels: 0

DB: 1 Gaps: 0

US-10-825-632-2 (1-3120) x FDL4W (1-85)

QY 26 GCCGCTGCTACTGCCGCCGCTGCT 49

DB 75 AlaAlaAlaThrAlaAlaAlaAla 82

RESULT 9

A22592

antifreeze protein IIA7 precursor - winter flounder

C:Species: Pseudopleuronectes americanus (winter flounder)

C:Date: 05-Oct-1988 #sequence_revision 05-Oct-1988 #text_change 09-Jul-2004

C:Accession: A22592; S12604

R:Gourlie, B.; Lin, Y.; Price, J.; De Vries, A.L.; Powers, D.; Huang, R.C.C.

J. Biol. Chem. 259, 14960-14965, 1984

A:Title: Winter flounder antifreeze proteins: a multigene family.

A:Reference number: A22592; MUID:85054993; PMID:6548752

A:Accession: A22592

A:Molecule type: mRNA

A:Residues: 1-91 <GOU>

A:CROSS-references: UNIPROT:P07835; UNIPARC:UPI0000125B4B; GB:M10148; NID:g213579; PIDN:1

R:Gauthier, S.; Wu, Y.; Davies, P.L.

Nucleic Acids Res. 18, 5303, 1990

A:Title: Nucleotide sequence of a variant antifreeze protein gene.

A:Reference number: S12604; MUID:90384854; PMID:2402466

A:Accession: S12604

A:Molecule type: DNA

A:Residues: 1-63, 'V', 65-91 <GAU>

A:CROSS-references: UNIPARC:UPI0000001382; EMBL:X53718; NID:g64213; PIDN:CAA37754.1; PID

C:Genetics:

A:Introns: 19/2

C:Superfamily: antifreeze protein

C:Keywords: antifreeze; tandem repeat

F:1-21/Domain: signal sequence #status predicted <SIG>

F:22-91/Product: antifreeze protein IIA7 #status predicted <MAT>

Alignment Scores:

Pred. No.: 81.3 Length: 91

Score: 8.00 Matches: 8

Percent Similarity: 100.0% Conservative: 0

Best Local Similarity: 100.0% Mismatches: 0

Query Match: 0.8% Indels: 0

DB: 2 Gaps: 0

US-10-825-632-2 (1-3120) x A22592 (1-91)

QY 26 GCCGCTGCTACTGCCGCCGCTGCT 49

DB 81 AlaAlaAlaThrAlaAlaAlaAla 89

```
RESULT 10
S02376
antifreeze protein precursor - yellowtail flounder
C:Species: limanda ferruginea (yellowtail flounder)
C>Date: 01-Dec-1989 #sequence_revision 01-Dec-1989 #text_change 09-Jul-2004
C:Accession: S02376
R:Scott, G.K.; Davies, P.L.; Shears, M.A.; Fletcher, G.L.
Eur. J. Biochem. 168, 629-633, 1987
A:Title: Structural variations in the alanine-rich antifreeze proteins of the Pleuronectes
A:Reference number: S02376; MUID:88029483; PMID:3665937
A:Accession: S02376
A:Molecule type: mRNA
A:Residues: 1-97 <CO>
A:Cross-references: UNIPROT:P09031; UNIPARC:UPI0000125B4D; EMBL:X06356; NID:964041; PIDN
F:14-48/Domain: propeptide #status predicted <PRO>
A:Note: part of this sequence, including the amino end of the mature protein, was confir
C:Superfamily: antifreeze protein
C:Keywords: antifreeze
F:123/Domain: signal sequence #status predicted <SIG>
F:14-48/Domain: propeptide #status predicted <PRO>
F:19-36/Product: antifreeze protein #status predicted <MAT>
Alignment Scores:
Pred. No.: 80.5 Length: 97
Score: 8.00 Matches: 8
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 0.8% Indels: 0
DB: 2 Gaps: 0
US-10-825-632-2 (1-3120) x S02376 (1-97)
QY 26 GCCGCTGCTACTGCCGCCGCTGCT 49
Db 58 AlaAlaAlaThrAlaAlaAlaAla 65
RESULT 11
B75632
hypothetical protein - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: B75632
R:White, O.; Elsen, J.A.; Heideberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: B75632; MUID:20036896; PMID:10567266
A:Accession: B75632
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-113 <WH>
A:Cross-references: UNIPROT:Q9RZK0; UNIPARC:UPI00000C15DC; GB:AE001826; NID:96460827; PI
A:Experimental source: strain R1
C:Genetics:
A:Gene: DRB0127
A:Map position: megaplasmid
A:Genome: plasmid
A:Note: plasmid MPI
C:Superfamily: Deinococcus radiodurans hypothetical protein DRB0127
Alignment Scores:
Pred. No.: 78.6 Length: 113
Score: 8.00 Matches: 8
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 0.8% Indels: 0
DB: 2 Gaps: 0
US-10-825-632-2 (1-3120) x B75632 (1-113)
QY 1367 CGAGTCTGGAGCGCATCTAGTAGG 1344
Db 17 ArgValTrpGluArgSerSerArg 24
```

```
RESULT 12
F64083
ribosomal protein L12 - Haemophilus influenzae (strain Rd KW20)
C:Species: Haemophilus influenzae
C>Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 09-Jul-2004
C:Accession: F64083
R:Gieschmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.
; Gokayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.
; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.
Science 269, 496-512, 1995
A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,
A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A:Reference number: A64000; MUID:95350630; PMID:7542800
A:Accession: F64083
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-123 <TIGR>
A:Cross-references: UNIPROT:P44348; UNIPARC:UPI00001685BA; GB:U32747; GB:L42023; NID:915
C:Superfamily: Escherichia coli ribosomal protein L12
C:Keywords: protein biosynthesis; ribosome
Alignment Scores:
Pred. No.: 77.5 Length: 123
Score: 8.00 Matches: 8
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 0.8% Indels: 0
DB: 2 Gaps: 0
US-10-825-632-2 (1-3120) x F64083 (1-123)
QY 48 GCAGCGGGCGAGTAGCAGCGGCC 25
Db 35 AlaAlaAlaValAlaAlaAla 42
RESULT 13
SFPQ
secretin precursor - pig
C:Species: Sus scrofa domestica (domestic pig)
C>Date: 24-Apr-1994 #sequence_revision 12-Apr-1996 #text_change 09-Jul-2004
C:Accession: B35094; A01544; A36052
R:Kopin, A.S.; Wheeler, M.B.; Leiter, A.B.
Proc. Natl. Acad. Sci. U.S.A. 87, 2299-2303, 1990
A:Title: Secretin: structure of the precursor and tissue distribution of the mRNA.
A:Reference number: A35094; MUID:90192795; PMID:2315322
A:Accession: B35094
A:Molecule type: mRNA
A:Residues: 1-131 <KOP>
A:Cross-references: UNIPROT:P01279; UNIPARC:UPI00001357C8; GB:M31496; NID:91646670; PIDN:
R:Mutt, V.; Jorpes, J.E.; Magnusson, S.
Eur. J. Biochem. 15, 513-519, 1970
A:Title: Structure of porcine secretin. The amino acid sequence.
A:Reference number: A91147; MUID:70282334; PMID:5465996
A:Accession: A01544
A:Molecule type: protein
A:Residues: 30-56 <MUT>
A:Cross-references: UNIPARC:UPI000002BEED
A:Note: tryptic peptides were sequenced
R:Gafvelin, G.; Joernvall, H.; Mutt, V.
Proc. Natl. Acad. Sci. U.S.A. 87, 6781-6785, 1990
A:Title: Processing of prosecretin: isolation of a secretin precursor from porcine intest
A:Reference number: A36052; MUID:90370867; PMID:2395872
A:Accession: A36052
A:Status: preliminary
A:Molecule type: protein
A:Residues: 30-59, 'R', 92-131 <GAP>
A:Cross-references: UNIPARC:UPI000017350D
R:Bohanszky, M.; Ondetti, M.A.; Levine, S.D.; Narayanan, V.L.; Saltza, M.V.; Sheehan, J.T.
Chem. Ind. 1966, 1757-1758, 1966
A:Title: Synthesis of a heptacosapeptide amide with the hormonal activity of secretin.
A:Reference number: A90916
A:Contents: annotation
```

A>Note: synthesis confirmed the proposed structure of the natural hormone

C:Superfamily: glucagon

C:Keywords: amidated carboxyl end; duodenal mucosa; duplication; hormone; secretagogue

F:1-18/Domain: signal sequence #status predicted <SIG>

F:30-56/Product: secretin #status experimental <MAT>

F:56/Modified site: amidated carboxyl end (Val) (amide in mature form from following gly

Alignment Scores:
Pred. No.: 76.8 Length: 131
Score: 8.00 Matches: 8
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 0.8% Indels: 0
DB: 1 Gaps: 0

US-10-825-632-2 (1-3120) x SEPG (1-131)

QY 30 CTGCTACTGCGCGCTGCTTCTT 53

Db 7 LeuLeuLeuProLeuLeuLeu 14

RESULT 14

F38164 hypothetical prtein 6 (cobo 3' region) - Pseudomonas sp.

C:Species: Pseudomonas sp.

C:Date: 31-Jul-1992 #sequence_revision 31-Jul-1992 #text_change 09-Dec-2002

C:Accession: F38164

R:Crouzet, J.; Levy-Schil, S.; Cameron, B.; Cauchois, L.; Rigault, S.; Rouyez, M.C.; Bla

J. Bacteriol. 173, 6074-6087, 1991

A:Title: Nucleotide sequence and genetic analysis of a 13.1-kilobase-pair Pseudomonas de

adenosyltransferase, cobyric acid synthase, and bifunctional cobinamide kinase-cobinami

A:Reference number: A38164; MUID:92011366; PMID:1655697

A:Accession: F38164

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-141 <CRO>

A:Cross-references: UNIPARC:UPI000013A59A; GB:M62866; NID:g151150; PIDN:AAA25782.1; PID:

C:Superfamily: Transporter DME family

Alignment Scores:

Pred. No.: 75.9 Length: 141
Score: 8.00 Matches: 8
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 0.8% Indels: 0
DB: 2 Gaps: 0

US-10-825-632-2 (1-3120) x F38164 (1-141)

QY 48 CTCTTAGTCCGCGTTCGCGCC 71

Db 10 LeuLeuSerAlaAlaPheAlaAla 17

RESULT 15

AE2962

conserved hypothetical protein Atu3299 [imported] - Agrobacterium tumefaciens (strain C5

C:Species: Agrobacterium tumefaciens

C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004

C:Accession: AE2962

R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I

erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayav, T.; Levy, R.; Li, M.; McClell

; Karp, P.; Romero, P.; Zhang, S.

Science 294, 2317-2323, 2001

A:Authors: Yoo, H.; Rao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,

ster, E.W.

A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A:Reference number: AB2577; MUID:21608550; PMID:11743193

A:Accession: AE2962

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-145 <KUR>

A:Cross-references: UNIPROT:Q8UAS1; UNIPARC:UPI00001647DE; GB:AE008689; PIDN:AAL44115.1;

A:Experimental source: strain C58 (Dupont)

C:Genetics:

A:Gene: Atu3299

A:Map position: linear chromosome

C:Superfamily: Transporter DME family

Alignment Scores:
Pred. No.: 75.5 Length: 145
Score: 8.00 Matches: 8
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 0.8% Indels: 0
DB: 2 Gaps: 0

US-10-825-632-2 (1-3120) x AE2962 (1-145)

QY 48 CTCTTAGTCCGCGTTCGCGCC 71

Db 13 LeuLeuSerAlaAlaPheAlaAla 20

Search completed: May 2, 2006, 03:51:57
Job time : 122 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: May 2, 2006, 03:11:14 ; Search time 137.5 Seconds
(without alignments)

3201.816 Million cell updates/sec

Title: US-10-825-632-2

Perfect score: 1030

Sequence: 1 agtgctaaagctccgagg.....agaattactaaaaaaa 3120

Scoring table

Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 60.0 , Fgapext 7.0
Delop 60.0 , Delext 60.0

Searched: 2166443 seqs, 705528306 residues

Post-processing

Total number of hits satisfying chosen parameters: 4332514

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DRV=xlp
-Q=/abs/ABSSWEB.spool/US10825632/runat_01052006_110235_4423/app_query.fasta_1
-DB=UniProt -QFMT=fastan -SUFFIX=oligo.n2p.rup -MINWATCH=0.1 -LOOPCL=0
-LOOPT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs02p
-USER=US10825632 @CNG 1 1 466 @runat 01052006 110235 4423 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG -DRV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPEXT=60 -XGAPEXT=60 -XGAPEXT=60 -XGAPEXT=60
-YGAPEXT=60 -YGAPEXT=60 -DELOP=60 -DELEXT=60

Database : UniProt 05.80.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	886	86.0	898	1	DPP8_HUMAN
2	186	18.1	892	1	DPP8_MOUSE
3	13	1.3	432	2	O75868_HUMAN
4	13	1.3	557	2	O5TXJ2_ANOGA
5	13	1.3	621	2	O7PTT8_ANOGA
6	13	1.3	740	2	O5TTK8_ANOGA
7	13	1.3	847	2	O6GR22_XENLA
8	13	1.3	862	1	O8BV94_mus musculus
9	13	1.3	863	1	DPP9_HUMAN
10	13	1.3	886	2	O7QBK1_ANOGA
11	13	1.3	923	2	O4SBM6_TETNG
12	13	1.3	1053	2	O9VC20_DROME
13	13	1.3	1113	2	O9VCL9_DROME
14	11	1.1	508	2	O75273_HUMAN
15	11	1.1	850	2	O6H9E3_9TRYP
16	11	1.1	852	2	O4Q1H9_LEIMA

17	11	1.1	1371	2	Q4P650_USFWA
18	10	1.0	289	2	Q4ZMT2_PSESI
19	10	1.0	478	2	Q9SP70_CARMA
20	10	1.0	564	2	Q8K4M2_RAT
21	10	1.0	893	2	Q9VK71_DROME
22	10	1.0	893	2	Q6NR25_DROME
23	9	0.9	87	2	Q8QFV4_LAMEL
24	9	0.9	103	2	Q9LJP7_ARATH
25	9	0.9	113	2	Q9LIW1_9ENTO
26	9	0.9	147	2	Q6D343_ERWCT
27	9	0.9	150	2	Q61CF4_CAEBR
28	9	0.9	154	1	PSBU_CVACA
29	9	0.9	155	2	Q5Z5HS_ORYSA
30	9	0.9	163	2	Q5Z4H9_ORYSA
31	9	0.9	172	2	Q8KAJ7_ORYSA
32	9	0.9	222	2	Q7N2V9_PHELL
33	9	0.9	230	2	Q5BW05_SCHJA
34	9	0.9	237	2	Q852K8_ORYSA
35	9	0.9	241	2	Q7NLG5_GLOVI
36	9	0.9	248	2	Q9PWL3_PETMA
37	9	0.9	249	1	PSA3_SFION
38	9	0.9	250	2	Q5YLZ5_ORYSA
39	9	0.9	256	2	Q4NVI6_9DEL
40	9	0.9	262	2	Q941I5_SORBI
41	9	0.9	265	2	Q8PHG7_XANAC
42	9	0.9	274	2	Q7P202_CHRVO
43	9	0.9	285	2	Q6N7X5_RHOPA
44	9	0.9	295	2	Q5EMP8_SORBI
45	9	0.9	297	2	Q9HST6_HALSA

ALIGNMENTS

RESULT 1
DPP8_HUMAN STANDARD; PRT: 898 AA.
ID DP8_HUMAN STANDARD; PRT: 898 AA.
AC Q6VLX1; Q7ZAC8; Q7ZAD3; Q7ZAE1; Q81WG7; Q8NEM5; Q96JX1; Q9HBM2;
AC Q9HBM3; Q9HBM4; Q9HBM5; Q9NKF4;
DT 13-SEP-2005 (Rel. 48, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Dipeptidyl peptidase 8 (EC 3.4.14.5) (Dipeptidyl peptidase VIII) (DP8)
DE (Prol) dipeptidase DPP8 (Dipeptidyl peptidase IV-related protein 1)
DE (DPP8-1).
GN Name=DPP8; Synonyms=DPRP1; ORFNames=MSTP097, MSTP135, MSTP141;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 3), NUCLEOTIDE SEQUENCE [MRNA] OF
RP 334-898 (ISOFORM 4), NUCLEOTIDE SEQUENCE [MRNA] OF 540-898 (ISOFORM
RP 5), NUCLEOTIDE SEQUENCE [MRNA] OF 260-792 (ISOFORM 6), FUNCTION,
RP CATALYTIC ACTIVITY. ENZYME REGULATION; TISSUE SPECIFICITY, INDUCTION,
RP AND SUBCELLULAR LOCATION.
RC TISSUE=Placenta;
RX MEDLINE=20467194; PubMed=11012666;
RA Abbott C.A., Yu D.M.T., Woollatt E., Sutherland G.R., McCaughan G.W.,
RA Gorrell M.D.;
RT "Cloning, expression and chromosomal localization of a novel human
RT dipeptidyl peptidase (DPP) IV homolog; DPP8.";
RL Eur. J. Biochem. 267:6140-6150(2000).
[2]
RN NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 3), CATALYTIC ACTIVITY, ENZYME
RP REGULATION, BIOPHYSICOCHEMICAL PROPERTIES, SUBCELLULAR LOCATION, AND
RP TISSUE SPECIFICITY.
RC TISSUE=Testis;
RX PubMed=12862155; DOI=10.1042/BJ20021914;
RA Qi S.Y., Riviere P.J., Trojnar J., Junien J.-L., Akinsanya K.O.;
RT "Cloning and characterization of dipeptidyl peptidase 10, a new member
RT of an emerging subgroup of serine proteases.";
RL Biochem. J. 373:179-185(2003).

RP NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 1).
RC TISSUE=Testis;
RA Sha J.H., Zhou Z.M., Li J.M.;
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
RN
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 4), AND NUCLEOTIDE
RC TISSUE=Hepatoma, and placenta;
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RW Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA Sekine M., Ohtsuka M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RW Yamamoto Y., Ito S., Saito K., Kawai Y., Isono Y., Nakamura Y.,
RA Watanabe K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M.,
RW Sugawara A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H.,
RA Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M.,
RW Yamazaki M., Ninomiya K., Ishibashi T., Yamashita H., Murakawa K.,
RA Ishida S., Ono Y., Takiguchi S., Watanabe M., Hiraoaka S., Chiba Y.,
RW Kusano Y., Kanehori K., Komai F., Hara R., Takeuchi K., Arita M.,
RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohara N., Aotoeuka S.,
RW Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakiuchi H.,
RW Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
RW Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujimura T.,
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,
RW Kawabata R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
RA Matsumura K., Nakajima T., Mizuno T., Morinaga M., Sasaki M.,
RW Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
RA Mitsuhashi-Sugano J., Satoh T., Shirai K., Takahashi Y., Nakagawa K.,
RW Okumura K., Nagase T., Nakamura Y., Ohara O., Isegai T., Sugano S.,
RA "Complete sequencing and characterization of 21,243 full-length human
RT cDNAs.";
RL Nat. Genet. 36:40-45(2004).
RN
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORMS 1 AND 2).
RC TISSUE=Testis;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RW Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RW Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Datchenko L., Matrusina K., Farmer A.A., Rubin G.M., Hong L.,
RW Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RW Beak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RW Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RW Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywicki M.I., Skalska U., Smalusz D.E.,
RW Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
RL and mouse cDNA sequences.";
RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 561-898.
RC TISSUE=Aorta;
RA Zhao B., Xu H.S., Tong Y.K., Sheng H., Qin B.M., Liu Y.Q., Liu B.,
RW Wang X.Y., Zhang Q., Song L., Gao Y., Zhang C.L., Ye J., Ji X.J.,
RA Liu B.H., Lu H., Chen J.Z., Cai M.Q., Zheng W.Y., Teng C.Y., Liu Q.,
RW Yu L.T., Lin J., Gong Q., Zhang A.M., Gao R.L., Hul R.T.;
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.

RP MUTAGENESIS OF GLU-275; SER-755; ASP-833 AND HIS-865, CATALYTIC
RC ACTIVITY, BIOPHYSICOCHEMICAL PROPERTIES, AND SUBCELLULAR LOCATION.
RA Ajami K., Abbott C.A., Obradovic M., Gysbers V., Kaehne T.,
RW McCaughan G.W., Gorrell M.D.;
RL in the CD36/DP1V gene family.";
RN Biochemistry 42:694-701(2003).
RN
RP BIOPHYSICOCHEMICAL PROPERTIES.
RC PubMed=15039077; DOI=10.1016/j.pep.2003.12.019;
RA Chen Y.-S., Chien C.-H., Goparaju C.M., Hsu J.T.-A., Liang P.-H.,
RW Chen X.;
RL "Purification and characterization of human prolyl dipeptidase DPP8 in
RT Sf9 insect cells.";
RN Protein Expr. Purif. 35:142-146(2004).
RN
RP ENZYME REGULATION.
RC PubMed=15664839; DOI=10.1016/j.bmcl.2004.11.023;
RA Jiang W.-T., Chen Y.-S., Hsu J., Wu S.-H., Chien C.-H., Chang C.-N.,
RW Chang S.-P., Lee S.-J., Chen X.;
RL "Novel isindoline compounds for potent and selective inhibition of
RT prolyl dipeptidase DPP8.";
RN Bioorg. Med. Chem. Lett. 15:687-691(2005).
RN
RP FUNCTION: Dipeptidyl peptidase that cleaves off N-terminal
RC dipeptides from proteins having a Pro or Ala residue at position
RW 2. May play a role in T-cell activation and immune function.
RL
CC CATALYTIC ACTIVITY: Release of an N-terminal dipeptide, Xaa-Yaa-|-
RW Zaa-, from a polypeptide, preferentially when Yaa is Pro, provided
CC Zaa is neither Pro nor hydroxyproline.
RW
CC ENZYME REGULATION: Inhibited by zinc. Inhibited by the serine
RW proteinase inhibitor 4-(2-aminoethyl)benzenesulphonyl fluoride
CC (AEBF), and by di-isopropylfluorophosphate. Specifically inhibited
RW by isindoline derivatives.
RN
RP BIOPHYSICOCHEMICAL PROPERTIES:
RC Kinetic parameters:
RW KM=208 uM for Ala-Pro-AMC;
KM=130 uM for Ala-Pro-AFC;
KM=120 uM for H-Ala-Pro-pNa;
KM=1420 uM for H-Ala-Ala-pNa;
KM=310 uM for H-Arg-Pro-pNa;
KM=2050 uM for H-Asp-Pro-pNa;
KM=480 uM for H-Gly-Pro-pNa;
pH dependence:
Optimum pH is 7.4-8.5. Little activity below pH 6.5;
SUBCELLULAR LOCATION: Cytoplasmic.
ALTERNATIVE PRODUCTS:
Event=Alternative splicing; Named isoforms=6;
Name=1;
IsoId=Q6V1X1-1; Sequence=Displayed;
Name=2;
IsoId=Q6V1X1-2; Sequence=VSP_013864;
Name=3;
IsoId=Q6V1X1-3; Sequence=VSP_013860;
Name=4;
IsoId=Q6V1X1-4; Sequence=VSP_013860, VSP_013862;
Name=5;
IsoId=Q6V1X1-5; Sequence=VSP_013863;
Name=6;
IsoId=Q6V1X1-6; Sequence=VSP_013861;
TISSUE SPECIFICITY: Ubiquitously expressed, with highest levels in
testis, placenta, prostate, muscle and brain.
INDUCTION: In activated T-cells.
SIMILARITY: Belongs to the peptidase S9B family. DPPIV subfamily.
CAUTION: Ref.4 (BAA91059) sequence differs from that shown due to
a frameshift in position 486.
CAUTION: Ref.6 (AAQ13650 and AAQ13623) sequences differ from that
shown due to several frameshifts.
This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its

CC use as long as its content is in no way modified and this statement is not
CC removed.

DR EMBL; AF221634; AAG29766.1; -; mRNA.
DR EMBL; AF221635; AAG29767.1; -; mRNA.
DR EMBL; AF221636; AAG29768.1; -; mRNA.
DR EMBL; AF221637; AAG29769.1; -; mRNA.
DR EMBL; AY172659; AA017261.1; -; mRNA.
DR EMBL; AY344202; AAQ63887.1; -; mRNA.
DR EMBL; AK000290; BAA91059.1; ALT_FRAME; mRNA.
DR EMBL; AK027826; BAB55395.1; ALT_INIT; mRNA.
DR EMBL; BC030688; AAH30688.3; -; mRNA.
DR EMBL; BC040203; AAH40203.1; ALT_INIT; mRNA.
DR EMBL; AF176779; AAQ13657.1; ALT_INIT; mRNA.
DR EMBL; AF175225; AAQ13650.1; ALT_FRAME; mRNA.
DR EMBL; AF173382; AAQ13623.1; ALT_FRAME; mRNA.

Alignment Scores:

Pred. No.:	0	Length:	898
Score:	886.00	Matches:	886
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	86.0%	Indels:	0
DB:	1	Gaps:	0

US-10-825-632-2 (1-3120) x DPP8_HUMAN (1-898)

QY	202	GGMAAATGCAATGGCAGCAGCAATGGAAACAGACAGCTGGGTTCAGATATTGAA	261
DB	13	GlyLysCysAsnMetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGlu	32
QY	262	ACTGGCGACTCTGAGGAGAATATTGAATCACAGATCGCGCTAAATTTGGAGCTTTTAT	321
DB	33	ThrAlaAspCysGluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyr	52
QY	322	GTTCAGCGGTATTCCTGGAGTCAGCTTAAAGAGCTGCTTGGCCGATACAGAAAAATATCAT	381
DB	53	ValGluArgTyrSerTrpSerGlnLeuLysLysLeuLeuAlaAspThrArgLysTyrHis	72
QY	382	GGCTACATGATGGCTAAGGCCACCATGATTTTCATGTTTGTGAAGAGAAATGATCCAGAT	441
DB	73	GlyTyrMetMetAlaLysAlaProHisAspPheValLysArgAsnAspProAsp	92
QY	442	GGACCTCATTCAGACAGATCTATTACCTTCGCATCTCTGGTGAGAACAGAGAAAATACA	501
DB	93	GlyProHisSerAspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThr	112
QY	502	CTGTTTATTCGAAATCCMAAACTATCAATAGACGACAGTCTTAATGCTCTCTTGG	561
DB	113	LeuPheTyrSerGluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrp	132
QY	562	AAGCTCTTTTGGATCTTTTTCAGGCAACTCGACTATGGAAATGATTCTCGAGAAGAA	621
DB	133	LysProLeuLeuAspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGlu	152
QY	622	GAACTATTAAAGAGAAAGAAACGCAATGGAACTCGGAATTCCTTTACGATTATCAC	681
DB	153	GluLeuLeuArgGluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHis	172
QY	682	CAAGGAATGGAAATTTCTGTGTTTCAAGCGGTAGTGGAAATTTATCATCGTAAAGATGGA	741
DB	173	GlnGlySerGlyThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGly	192
QY	742	GGGCCAAGGATTTAGCAACACCTTTAAGCCCAATCTAGTGGAACTAGTTGTCCC	801
DB	193	GlyProGlnGlyPheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysPro	212
QY	802	AACATACGGATGGATCCAAAATTATGCCCCGCTGATCCAGACTGGATTGCTTTATACAT	861
DB	213	AsnIleArgMetAspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHis	232
QY	862	AGCAACGATATTCGATATCTTAACATCGTAACAGAGAAAGGAGACTCATTTATGTG	921

DB	233	SerAsnAspIleTrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrVal	252
QY	922	CACATGAGCTAGCAACATCGAAGAGATGCCAGATCAGCTGGAGTCCCTACCTTGT	981
DB	253	HisAsnGluLeuAlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheVal	272
QY	982	CTCCAAGAAGAAATTTGATAGATATTCCTGCTATTGGTGTGTCAAAAGCTCAAACT	1041
DB	273	LeuGlnGluGluPheAspArgTyrSerGlyTyrTrpIleCysProLysAlaGluThrThr	292
QY	1042	CCCAGTGTGTGTAATAATCTTAGAATTTCTATATGAAAGAAATGATGAATCTGAGTGGAA	1101
DB	293	ProSerGlyGlyLysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGlu	312
QY	1102	ATTATTGATGTTATCATCCCTATGTTGGAAACAAAGAGGCGCAGATTCATTCGGTTATCCT	1161
DB	313	IleIleHisValThrSerProMetLeuGluThrArgAlaAspSerPheArgTyrPro	332
QY	1162	AAAACAGGTACAGCAAAATCCTAAAGTCACCTTTTAAAGTCTCAGAAATATGATGCT	1221
DB	333	LysThrGlyThrAlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAla	352
QY	1222	GAAAGAGGATCATAGATGTCATAGATAAGGAACTAATCAACCTTTTCAGATTCCTATT	1281
DB	353	GluGlyArgIleIleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPhe	372
QY	1282	GAAAGAGTGAATATATTGCGCAGCTGGATGACTCCTGAGGAAAAATATGCTGGTCC	1341
DB	373	GluGlyValGluTyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSer	392
QY	1342	ATCCTACTAGATCGCTCCCAGACTCGCCTACAGATAGTGTGATCTCACCTGAATTTT	1401
DB	393	IleLeuLeuAspArgSerGlnThrArgLeuGlnIleValIleLeuSerProGluLeuPhe	412
QY	1402	ATCCAGTGAAGATGATGTTGGAAGCGACAGACTCATTTGAGTCACTGCTGCTGATCT	1461
DB	413	IleProValGluAspAspValMetGluArgGlnArgLeuIleGluSerValProAspSer	432
QY	1462	GTGAGCCACATAATTTCTATGAAGAAACAAACAGACATCTGGATAAATATCCATGATC	1521
DB	433	ValThrProLeuIleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIle	452
QY	1522	TTTCATGTTTTTCCCAAGTCACGAGAGAAATGAGTTTATTTTTCCTGCTGAATGC	1581
DB	453	PheHisValPheProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCys	472
QY	1582	AAACACAGTTTCCCTCATTTATACAAATTTACATCTATTAAAGGAAAGCAAAATATAA	1641
DB	473	LysThrGlyPheArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLys	492
QY	1642	CGATCCAGTGTGGGCTCGCTCCCAAGTGAATTTCAAGTGTCTTCAAGAGGAGATA	1701
DB	493	ArgSerSerGlyGlyLeuProAlaProSerAspPheLysCysProIleLysGluIle	512
QY	1702	GCAATTACAGTGTGTAATGGGAAGTTCCTTGGCGCGGATGATCTTAATATCCAAAGTGA	1761
DB	513	AlaIleThrSerGlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAsp	532
QY	1762	GAACTCAGAGCGCTGATATTTTGAAGACCAAAAGACTCCCTTTTAGAGCATCACCTG	1821
DB	533	GluValArgLeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeu	552
QY	1822	TACGTAGTCAGTACGTAATAATCCTGGAGAGGTGCACAAAGCTGACTGACCGTGGCTACTCA	1881
DB	553	TyrValValSerTyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSer	572
QY	1882	CATTCTTCTGCTCATCAGTCAGCTGTGACTCTTTTATAAGTAAGTATAGTAAACAGAG	1941
DB	573	HisSerCysLysIleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLys	592
QY	1942	AATCCACACTGTGTGCTCCCTTTTCAAGCTATCAAGTCTCAAGATGACCACTTGCAAA	2001
DB	593	AsnProHisCysValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLys	612

QY 2002 ACRAAGCAATTTGGGCCACCATTTGGATTACGAGGTCCCTCTCTGACTACTACCT 2061
Db |||||||
QY 613 ThrlySGluPheThrPalaThrIleLeuApsSerAlaGlyProLeuProAspIleThrPro 632
QY 2062 CCAGAAATTTCTCTTTTGAAGACTACTCGGATTTACATTTGATGGGATGCTCTACAAG 2121
Db |||||||
QY 633 ProGluIlePheSerPheGluSerThrThrGlyPheThrLeuTyrglyMetLeuTyrllys 652
QY 2122 CCTCATGATCTACAGCTCGAAGAAAGAAATATCTCTACTGTGCTCTTCATATATGGTGGTCT 2181
Db |||||||
QY 653 ProHisAspLeuGlnProGlyLysLysTyProThrValLeuPheIleTyrglyGlyPro 672
QY 2182 CAGGTGCGATTTGGTGAATAATTCGTTTAAAGGAGTCAAGTATTTCCGGTTTCAATACCCCTA 2241
Db |||||||
QY 673 GlnValGlnLeuValAsnAsnArgPheLysGlyValLysTyPheArgLeuAsnThrLeu 692
QY 2242 GCCTCTCTAGGTATGTGGTTGTAGTGATAGACAACAGGGGATCTCTGTCACCGAGGGCTT 2301
Db |||||||
QY 693 AlaSerLeuGlyTyValValValValIleAspAsnArgGlySerCysHisArgGlyLeu 712
QY 2302 AAATTTGAAGCGCTTTTAAATATAAATGGGTCAAAATAGAAATTCAGCATCAGGTGGAA 2361
Db |||||||
QY 713 LysPheGluGlyAlaPheLysTyLysMetGlyGlnIleGluIleAspAspGlnValGlu 732
QY 2362 GGACTCCAATATCTAGCTCTCCGATATGATTTTCATGACTTAGCTGCTGGCATCCAC 2421
Db |||||||
QY 733 GlyLeuGlnTyLeuAlaSerArgTyArgPheIleAspPueAspArgValGlyIleHis 752
QY 2422 GCCTGGTCTTGGAGGATACCTCTCCGATGGCATTAATCCAGAGTCAGATATCTTC 2481
Db |||||||
QY 753 GlyTrpSerTyrglyGlyTyLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePhe 772
QY 2482 AGGTTGCTATGCTGGGGCCCGACCTCTCTGTGTGATCTCTATGATACAGGATACAG 2541
Db |||||||
QY 773 ArgValAlaIleAlaGlyAlaProValThrLeuTrpIlePheTyArgThrGlyTyThr 792
QY 2542 GAACGTTATATGGGTCACTCCCTGACCAAGATGAACAGGCTATTAATCTAGGATCTGTGGCC 2601
Db |||||||
QY 793 GluArgTyMetGlyHisProAspGlnAsnGluGlnGlyTyTyLeuGlySerValAla 812
QY 2602 ATCAAGCAGAAAGTTCCCTCTCGAACCAAAATCGTTTACTGCTCTTACATGGTTTCTCG 2661
Db |||||||
QY 813 MetGlnAlaGlyLysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeu 832
QY 2662 GATGAGATGTCCATTTTGACATACACAGTATATTAATCTAGTGGTGGTGGGCTGGA 2721
Db |||||||
QY 833 AspGluAsnValHisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGly 852
QY 2722 AAGCATATGATTTTACAGATCTATCTCTCAGGAGACACAGCATAAAGTTTCTCGAATCG 2781
Db |||||||
QY 853 LysProTyArgLeuGlnIleTyProGlnGluArgHisSerIleArgValProGluSer 872
QY 2782 GGAGACATATGAAGTGCATCTTTTGCACCTACCTCAAGAAACCTTGGATCAGTATT 2841
Db |||||||
QY 873 GlyGluHisTyGluLeuHisLeuLeuHisTyLeuGlnGluAsnLeuGlySerArgIle 892
QY 2842 GCTGCTCTAAAGTGATA 2859
Db |||||||
QY 893 AlaAlaLeuLysValIle 898

RESULT 2

DPP8_MOUSE
ID DPP8_MOUSE STANDARD; PRT; 892 AA.
AC Q80YA7; Q80YA6;
DT 13-SEP-2005 (Rel. 48, Created)
DT 13-SEP-2005 (Rel. 48, Last sequence update)
DE 13-SEP-2005 (Rel. 48, Last annotation update)
DE Dipeptidyl peptidase 8 (EC 3.4.14.5) (Dipeptidyl peptidase VIII)
DE (Df8).
GN Name=Dpp8;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
Murioidea; Muridae; Murinae; Mus.
NCBI_taxID=10090;
[1]
NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nikaide I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Brusci V., Chothia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawai H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Kongsava A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Mik H.,
RA Nagashima T., Numa K., Okido T., Paven W.J., Pertea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Sample C.A., Setou M., Shimada K.,
RA Sulemana R., Takanaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wyshaw-Portis A., Yanagisawa M., Yang L.,
RA Yuan Z., Zavalan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa I., Kono H., Nakamura M., Sakazume N., Sato K.,
RA Shiraiki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.,
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573 (2002).
[2]
NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC STRAIN=C57BL/6J; TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.A., McWeeny P.J., McKernan K.J., Abramson R.D., Mullaly S.J.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettaman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butcherfield Y.S.N., Krzywinski M.I., Skalska U., Smalium D.E.,
RA Scherch A., Schein J.B., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
[3]
PHOSPHORYLATION SITE TYR-325.
RX PubMed=14729942; DOI=10.1074/mcp.D300003-MCP200;
RA Shu H., Chen S., Bi Q., Mumby M., Breken D.L.;
RT "Identification of phosphoproteins and their phosphorylation sites in
RT the WEHI-231 B lymphoma cell line."
RL Mol. Cell. Proteomics 3:279-286 (2004).
CC -!- FUNCTION: Dipeptidyl peptidase that cleaves off N-terminal
CC dipeptides from proteins having a Pro or Ala residue at position
CC 2. May play a role in T-cell activation and immune function (By
CC similarity).
CC -!- CATALYTIC ACTIVITY: Release of an N-terminal dipeptide, Xaa-Yaa-|-
CC Zaa-, from a polypeptide, preferentially when Yaa is Pro, provided

CC Zaa is neither Pro nor hydroxyproline.
CC -1- ENZYME REGULATION: Inhibited by zinc. Inhibited by the serine
CC proteinase inhibitor 4-(2-aminoethyl)benzenesulphonyl fluoride
CC (AEB5F), and by di-isopropylfluorophosphate. Specifically inhibited
CC by isindoline derivatives (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the peptidase S9B family. DPPIV subfamily.
CC -----
CC This Swiss-Prot entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; BAB30295.2; -; mRNA.
CC EMBL; BC043124; AAH43124.1; -; mRNA.
CC EMBL; BC059222; AAH59222.1; -; mRNA.
CC MEROPS; S09.018; -.
CC Ensembl; ENSMUSG00000032393; Mus musculus.
CC MGI; MGI:1921638; Dpp8.
CC InterPro; IPR001375; Peptidase_S9.
CC InterPro; IPR002469; Peptidase_S9B.
CC InterPro; IPR000379; Ser_estrs.
CC Pfam; PF00930; DPPIV_N; 1.
CC Pfam; PF00326; Peptidase_S9; 1.
CC Amino-peptidase; Hydrolase; Phosphorylation; Protease; Serine protease.
FT ACT_SITE 749 749 Charge relay system (By similarity).
FT ACT_SITE 827 827 Charge relay system (By similarity).
FT ACT_SITE 859 859 Charge relay system (By similarity).
FT MOD_RES 325 325 Phosphotyrosine.
FT CONFLICT 87 87 G -> R (in Ref. 1).
SQ SEQUENCE 892 AA; 102186 MW; 59081CD9792E03ED CRC64;

Alignment Scores:
Pred. No.: 1.39e-184 Length: 892
Score: 186.00 Matches: 286
Percent Similarity: 99.7% Conservative: 0
Best Local Similarity: 99.7% Mismatches: 1
Query Match: 18.1% Indels: 0
DB: 1 Gaps: 0

US-10-825-632-2 (1-3120) x DPP8_MOUSE (1-892)

QY 1999 AACAAGAAATTTGGGCCACCATTTGGATTGAGATTGAGGAGTCTCTCTGACTACT 2058
DB 606 LysThrLysGluPheTrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThr 625

QY 2059 CTTCCAGAAATTTCTCTTTGAAAGTACTGCTGATTTACATTTGATGGGATCTCTAC 2118
DB 626 ProProGluIlePheSerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyr 645

QY 2119 AAGCCTCATGCTACAGCTCGAAGAAATATCTCTGCTGCTTCATATATGTTGGT 2178
DB 646 LysProHisAspLeuGlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGly 665

QY 2179 CTTCCAGGTGAGTGTGTAATAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACC 2238
DB 666 ProGlnValGlnLeuValAsnAsnA-gPheLysGlyValLysTyrPheArgLeuAsnThr 685

QY 2239 CTAGCCTCTCTAGGTTATGTTGTTAGTGATAGACACAGGGGATCTCTGTCACCGAGGG 2298
DB 686 LeuAlaSerLeuGlyTyrValValValValIleAspAsnArgGlySerCysHisArgGly 705

QY 2299 CTTAAATTTGAGCGCTTTAAATATAAATGGGTCAATAGAAATTCAGATCAGGTG 2358
DB 706 LeuLysPheGluGlyAlaPheLysTyrLysMetGlyGlnIleGluIleAspGlnVal 725

QY 2359 GAAGGACTCCATATCTAGCTTCTCGATATGATTTTCATTTGACTGATCGTGTGGCATC 2418
DB 726 GluGlyLeuGlnTyrLeuAlaSerGlnTyrAspPheIleAspLeuAspArgValGlyIle 745

QY 2419 CACGCTGGTCTTATGAGGATACCTCTCCCTGATGGCAATTAATGACAGGTTCAGATATC 2478

DB 746 HisGlyTrpSerTyrGlyTyrGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIle 765
QY 2479 TTCAGGTTGCTATTGCTGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGATAC 2538
DB 766 PheArgValAlaIleAlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyr 785
QY 2539 ACGGAACGTTATATGGTGCACCTCGACCAATGAACAGCGCTATTACTTAGGATCTGTG 2598
DB 786 ThrGluArgTyrMetGlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerVal 805
QY 2599 GCCATCAAGCAGAAAGTTCCTCTCTGAAACCAATGTTTACTGCTCTTACATGGTTTC 2658
DB 806 AlaMetGlnAlaGluLysPheProSerGluProAsnArgLeuLeuLeuHisGlyPhe 825
QY 2659 CTGATGAGAAATGTCCTTTTCACATACCATGATATATCTAGCTTTTGTAGTGAGGCT 2718
DB 826 LeuAspGluAsnValHisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAla 845
QY 2719 GGAAGCCATATGATTACAGATCTATCTCAGGAGACACACAGATAAGAGTTCTCTGAA 2778
DB 846 GlyLysProTyrAspLeuGlnIleTyrProGlnGluArgHisSerIleArgValProGlu 865
QY 2779 TCGGAGAACATTTATGAACTGCTCTTTGCACTACTCTTCAAGAAACCTTGGATCAGT 2838
DB 866 SerGlyGluHisTyrGluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArg 885
QY 2839 ATTGCTGCTCTAAAGTGATA 2859
DB 886 IleAlaAlaLeuLysValIle 892

RESULT 3
ID 075868 HUMAN PRELIMINARY; PRT; 432 AA.
AC 075868;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DE 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE R33083.1 (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE
RA Lamerdin J.E., McCready P.M., Skowronski E., Viswanathan V.,
RA Burkhardt-Schultz K., Gordon L., Dias J., Ramirez M., Stilwagen S.,
RA Phan H., Velasco N., Do L., Regala W., Terry A., Ganes J.,
RA Dangnan L., Erler A., Christensen M., Georgescu A., Avila J., Liu S.,
RA Attix C., Andreise T., Trankheim M., Amico-Keller G., Coesfield J.,
RA Duarte S., Lucas S., Bruce R., Thomas P., Quan G., Kronmiller B.,
RA Arellano A., Saunders C., Ow D., Nolan M., Trong S., Kobayashi A.,
RA Olsen A.S., Carraro A.V.;
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC005783; AAC62840.1; -; Genomic DNA.
DR Ensembl; ENSG00000142002; Homo sapiens.
DR InterPro; IPR002469; Peptidase_S9B.
DR Pfam; PF00930; DPPIV_N; 1.
FT NON TER 432 432
SQ SEQUENCE 432 AA; 48595 MW; 64E2B88BE0523A7E CRC64;

Alignment Scores:
Pred. No.: 0.00255 Length: 432
Score: 13.00 Matches: 13
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 1.3% Indels: 0
DB: 2 Gaps: 0

US-10-825-632-2 (1-3120) x 075868_HUMAN (1-432)

QY 607 TATTCTGAGAGAGAACTATTATGAGAGAAACGC 645

DR InterPro; IPR002469; Peptidase_S9B.
DR GO; GO:000379; Ser. esters.
DR Pfam; PF00930; DPPIV_N; 1.
DR Pfam; PF00326; Peptidase_S9; 1.
FT NON TER 1
SQ SEQUENCE 740 AA; 84151 MW; D59F131BBA095CCB CRC64;

Alignment Scores:
Pred. No.: 0.00243 Length: 740
Score: 13.00 Matches: 13
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 1.3% Indels: 0
Gaps: 0
DB:

US-10-825-632-2 (1-3120) x Q5TTK8_ANOGA (1-740)

Qy 2416 ATCCACGGTGGTCTATGAGGATACCTCTCCCTGATG 2454
|||||
Db 603 ILeHisGlyTrpSerTyrGlyGlyTyrLeuSerLeuMet 615
|||||

RESULT 7
Q6GR22_XENLA
ID Q6GR22_XENLA PRELIMINARY; PRT; 847 AA.
AC Q6GR22;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE MGC81313 protein.
GN Name=MGC81313;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Embryo;
RX MEDLINE=2238257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner K.H., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grinwood J., Schmutz J., Myers R.M.,
RA Butterfield V.S.N., Kryzyski M.I., Skalska U., Smalish D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Embryo;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
initiative.";
RT Dev. Dyn. 225:384-391 (2002).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Embryo;
RX Klein S., Gerhard D.S.;
RT Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC071112; AAH71112.1; -, mRNA.
DR GO; GO:0003824; F.catalytic activity; IEA.
DR GO; GO:0008236; F.serine-type peptidase activity; IEA.
DR GO; GO:0006508; P.proteolysis and peptidolysis; IEA.
DR InterPro; IPR001375; Peptidase_S9.
DR InterPro; IPR002469; Peptidase_S9B.
DR InterPro; IPR000379; Ser. esters.
DR Pfam; PF00930; DPPIV_N; 1.
DR Pfam; PF00326; Peptidase_S9; 1.
SQ SEQUENCE 847 AA; 97484 MW; F515F2609B267BAF CRC64;

Alignment Scores:
Pred. No.: 0.0024 Length: 847
Score: 13.00 Matches: 13
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 1.3% Indels: 0
Gaps: 0
DB:

US-10-825-632-2 (1-3120) x Q6GR22_XENLA (1-847)

Qy 607 TATTCTCGAGAGAGAACTATTAAAGAGAAAGAAACGC 645
|||||
Db 109 TySerArgGluGluLeuLeuArgGluArgLysArg 121
|||||

RESULT 8
DPP9_MOUSE
ID DPP9_MOUSE STANDARD; PRT; 862 AA.
AC Q8BVG4; G6KAM9; Q8BWT9;
DT 13-SEP-2005 (Rel. 48, Created)
DT 13-SEP-2005 (Rel. 48, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Dipeptidyl peptidase 9 (EC 3.4.14.5) (Dipeptidyl peptidase IX) (DPP9)
DE (Dipeptidyl peptidase-like protein 9) (DPLP9).
GN Name=Dpp9;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 1).
RC STRAIN=C57BL/6J; TISSUE=liver, and Olfactory bulb;
RC MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
RX Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogi A., Schonbach C., Gojobori T.,
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Brad D., Brusci V., Chothia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasawa Y., Kedierski R.M., King B.L.,
RA Kongaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sadelin A., Schneider C., Sample C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Valverde R., Wagner L., Wahlestedt C., Wang Y., Wanabe Y., Wells C.,
RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shingawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).

RT	cytoplasmic localization and DPV-like peptidase activity.";	CC	Zaa-, from a polypeptide, preferentially when Yaa is Pro, provided
RL	Biochim. Biophys. Acta 1679:18-28(2004).	CC	Zaa is neither pro nor hydroxyproline.
RN	[4]	CC	-!- ENZYME REGULATION: Inhibited by the serine proteinase inhibitor 4-
RP	NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 1).	CC	(2-aminoethyl)benzenesulphonyl fluoride (AEBSF), and by di-
RX	TISSUE=Placenta, and Skin;	CC	isopropylfluorophosphate.
RA	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;	CC	Kinetic parameters:
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,	CC	KM=161 uM for Ala-Pro-AMC;
RA	Klausner R.D., Collins F.S., Wegner L., Shemen C.M., Schuler G.D.,	CC	KM=180 uM for Ala-Pro-AFC;
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,	CC	pH dependence:
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J.M., Hsieh F.,	CC	Optimum pH is 7.5-8.5. Little activity below pH 6.5;
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,	CC	-!- SUBCELLULAR LOCATION: Cytoplasmic; cytosol.
RA	Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,	CC	-!- ALTERNATIVE PRODUCTS:
RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,	CC	Event-Alternative splicing; Named isoforms=5;
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullighan S.J.,	CC	Name=1; Synonyms=Short;
RA	Boak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,	CC	Isoid=Q86T12-1; Sequence=Displayed;
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,	CC	Name=2; Synonyms=Long;
RA	Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,	CC	Isoid=Q86T12-2; Sequence=VSP_013865;
RA	Faney J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,	CC	Note=Incomplete sequence;
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,	CC	Name=3;
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,	CC	Isoid=Q86T12-3; Sequence=VSP_013867, VSP_013868;
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,	CC	Name=4;
RA	Butterfield V.S.N., Krzywinski M.I., Skalska U., Smallick D.E.,	CC	Isoid=Q86T12-4; Sequence=VSP_013869;
RA	Schmerch A., Schein J.B., Jones S.J.M., Marra M.A.,	CC	Name=5;
RT	"Generation and initial analysis of more than 15,000 full-length human	CC	Isoid=Q86T12-5; Sequence=VSP_013865, VSP_013866;
RT	and mouse cDNA sequences.";	CC	Note=Incomplete sequence;
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).	CC	-!- TISSUE SPECIFICITY: Ubiquitously expressed, with highest levels in
RN	[5]	CC	liver, heart and muscle, and lowest levels in brain.
RP	NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 30-863 (ISOFORM 3).	CC	-!- SIMILARITY: Belongs to the peptidase S9B family. DPVIV subfamily.
RP	NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 272-863 (ISOFORM 2), AND	CC	-!- CAUTION: Ref.6 (CAD39039) sequence differs from that shown due to
RP	PARTIAL NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 5).	CC	frameshifts in positions 432 and 460.
RC	TISSUE=Glial tumor, Ovary, Spleen, and Trachea;	CC	-----
RX	PubMed=14702039; DOI=10.1038/ng1285;	CC	This Swiss-Prot entry is copyright. It is produced through a collaboration
RA	Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,	CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
RA	Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,	CC	the European Bioinformatics Institute. There are no restrictions on its
RA	Sekine M., Ohayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,	CC	use as long as its content is in no way modified and this statement is not
RA	Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,	CC	removed.
RA	Nagahari K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M.,	CC	-----
RA	Shiratori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H.,	DR	EMBL; AF452102; AAL47179.1; -; mRNA.
RA	Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E.,	DR	EMBL; AY12660; AAO17362.1; -; mRNA.
RA	Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M.,	DR	EMBL; AF542510; AAO73880.2; -; mRNA.
RA	Yamazaki M., Ninomiya K., Ishibashi T., Yamashita H., Murakawa K.,	DR	EMBL; AY374518; AAO83119.1; -; mRNA.
RA	Fujimori K., Tanai H., Kinata M., Watanabe M., Hiraoka S., Chiba Y.,	DR	EMBL; BC000970; AAH00970.1; -; mRNA.
RA	Ishida S., Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T.,	DR	EMBL; BC037948; AAH37948.1; -; mRNA.
RA	Kusano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,	DR	EMBL; AK054656; BAB70784.1; ALT INIT; mRNA.
RA	Numura Y., Togiya S., Komai F., Hara R., Takeuchi K., Arita M.,	DR	EMBL; AK075030; BAC11362.1; -; mRNA.
RA	Imose N., Mueshino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,	DR	EMBL; AK131499; BADI8643.1; ALT INIT; mRNA.
RA	Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,	DR	EMBL; AL834376; CAD39039.3; ALT_FRAME; mRNA.
RA	Moriya S., Moniyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,	DR	EMBL; CR627380; CAH10477.1; -; mRNA.
RA	Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,	DR	MEROFS; S09.019; -; -
RA	Hishigaki H., Watanabe K., Kumagai A., Takemoto M., Kawakami B.,	DR	Ensembl; ENSG00000142002; Homo sapiens.
RA	Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,	DR	HGNC; HGNC:18648; DPP9.
RA	Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,	DR	MIM; 608258; -
RA	Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,	DR	InterPro; IPR001375; Peptidase_S9.
RA	Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,	DR	InterPro; IPR002469; Peptidase_S9B.
RA	Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,	DR	InterPro; IPR000379; Ser.estr.
RA	Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,	DR	Pfam; PF00930; DPVIV N; 1.
RA	Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,	DR	Pfam; PF00326; Peptidase_S9; 1.
RA	Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,	KW	Alternative splicing; Amino peptidase; Hydrolase; Protease;
RA	Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,	KW	Serine protease.
RA	Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;	FT	ACT SITE 730 730 Charge relay system (By similarity).
RT	"Complete sequencing and characterization of 21,243 full-length human	FT	ACT SITE 808 808 Charge relay system (By similarity).
RT	cDNAs.";	FT	ACT SITE 840 840 Charge relay system (By similarity).
RL	Nat. Genet. 36:40-45(2004).	FT	VARSPLIC 1 1 M -> LSRVPCVRCRPPPLPGLPGSQSRMRDREAPLD
RN	[6]	FT	PGPACGRPTSRVSUSHACSWNGGSLDPLECTPALLRSAE
RP	NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 209-863 (ISOFORM 4), AND	FT	RLMRKVKRLDKENTGWSRWSFLNSEGAER (in
RP	NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 298-863 (ISOFORM 2).	FT	isoform 2 and isoform 5).
RC	TISSUE=Melanoma;	FT	/FTId=VSP_013865.
RG	The German cDNA consortium;	FT	Missing (in isoform 5).
RL	Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.	FT	/FTId=VSP_013866.
CC	-!- FUNCTION: Dipeptidyl peptidase that cleaves off N-terminal	FT	QLVNNSPFGIKYRLRLNLASLGAV -> SAHLPRPPPHH
CC	dipeptides from proteins having a Pro or Ala residue at position	FT	PPDSPSPFLKQL (in isoform 3).
CC	2.		
CC	-!- CATALYTIC ACTIVITY: Release of an N-terminal dipeptide, Xaa-Yaa- -		

```
FT VARSPLIC 675 863 /FtId=VSP_013867.
FT Missing (in isoform 3).
FT /FtId=VSP_013868.
FT Missing (in isoform 4).
FT /FtId=VSP_013869.
FT I -> N (in Ref. 3).
FT CONFLICT 204 204
FT CONFLICT 571 571 C -> W (in Ref. 5; BAC85150).
FT CONFLICT 709 709 L -> P (in Ref. 5; BAB18643).
FT CONFLICT 753 753 G -> C (in Ref. 5; BAB70784).
SQ SEQUENCE 863 AA; 98263 MW; 40FE0B78E26CDED5 CRC64;

Alignment Scores:
Pred. No.: 0.0024 Length: 863
Score: 13.00 Matches: 13
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 1.3% Indels: 0
DB: 1 Gaps: 0

US-10-825-632-2 (1-3120) x DPP9_HUMAN (1-863)
QY 607 TATTCTCGAAGAAAGAACTATTAAAGAAAGAAACGC 645
Db 121 TyrSerArgGluGluLeuArgGluArgLysArg 133

RESULT 10
Q7QBK1 ANOGA
ID Q7QBK1 ANOGA PRELIMINARY; PRT; 886 AA.
AC Q7QBK1;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DE ENSANGP0000016526 (Fragment).
GN ORFNames=ENSANGG0000014037;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Culicidae;
OC Anophelinae; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
preliminary data.
DR EMBL; AAB01008979; EAA08416.2; -; Genomic_DNA.
DR MEROPS; S09.016; -.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001375; Peptidase_S9.
DR InterPro; IPR002469; Peptidase_S9B.
DR InterPro; IPR000379; Ser_estra.
DR Pfam; PF00930; DPPIV N; I.
DR Pfam; PF00326; Peptidase_S9; 1.
FT NON TER 1
FT NON TER 886
SQ SEQUENCE 886 AA; 100440 MW; 3C284605CAA57DB4 CRC64;

Alignment Scores:
Pred. No.: 0.00239 Length: 886
Score: 13.00 Matches: 13
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 1.3% Indels: 0
```

```
DB: 2 Gaps: 0
US-10-825-632-2 (1-3120) x Q7QBK1 ANOGA (1-886)
QY 2416 ATCCACGGCTGCTCTATGAGGATACCTCTCCCTCATG 2454
Db 750 ILEHISGLYRPSERTYGLYGLYTYRLEUSERLEUMET 762

RESULT 11
Q4SBM6_TETNG
ID Q4SBM6_TETNG PRELIMINARY; PRT; 923 AA.
AC Q4SBM6;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Chromosome 15 SCAF14667, whole genome shotgun sequence.
DE (Fragment).
GN ORFNames=GSTENG0020903001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Jallion O., Aury J.M., Brunet F., Petit J.L., Strange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dessat C., Segurens B.,
RA Anhouart V., Salenoubat M., Levy M., Baudet N., Castellano S.,
RA Bismont C., Skalli Z., Cattolico L., Katinka M., Vacherie B.,
RA Craud C., Duprat S., Brottier P., Coutanceau J.P., Guzy J.,
RA Parizot M., Lardier G., Chappie C., McKernan K.J., McEwan P., Bosak S.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissenbach J., Roest Crolius H.;
RT Genome duplication in the teleost fish Tetraodon nigroviridis reveals
the early vertebrate proto-karyotype."
RL Nature 431:946-957(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
preliminary data.
DR EMBL; CAAE01014667; CAG01956.1; -; Genomic_DNA.
FT NON TER 1
FT NON TER 923
SQ SEQUENCE 923 AA; 105211 MW; 1397023B2004D009 CRC64;

Alignment Scores:
Pred. No.: 0.00239 Length: 923
Score: 13.00 Matches: 13
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 1.3% Indels: 0
DB: 2 Gaps: 0

US-10-825-632-2 (1-3120) x Q4SBM6_TETNG (1-923)
QY 607 TATTCTCGAAGAAAGAACTATTAAAGAAAGAAACGC 645
Db 112 TyrSerArgGluGluLeuArgGluArgLysArg 124

RESULT 12
Q9VC20_DROME
ID Q9VC20_DROME PRELIMINARY; PRT; 1053 AA.
AC Q9VC20; Q8IH07;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
```


GN Name=CG3744; ORFNames=CG3744;
OS Drosophila melanogaster (Fruit fly)
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Phyllophaga; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Anantades P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers J.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abrell J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burris K.C., Buesam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harslin D., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hosteln D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpow G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laeko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattai B., McIntosh T.C., McLeod M.P., McPherson P.,
RA Merkulov G., Milshin N.V., Mobarry L., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusslein D.R., Pacle J., Reese M.G.,
RA Palazzolo M., Patton G.S., Fan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Staden-Klamon I., Simpson M.C., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskaas R., Tector A.C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.N., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RN Science 287:2185-2195(2000).
[2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426065; PubMed=12537568;
RA Ceiniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Friese E., Hodgson A.,
RA George R.A., Hoskins R.A., Laverty T., Muzny D.M., Nelson C.R.,
RA Pacle J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA Svirskaas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weinstein G., Scher S.E., Myers E.W., Gibbs R.A., Rubin G.M.,
RT "Finishing a whole-genome shotgun: release 3 of the drosophila
melanogaster euchromatic genome sequence";
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
[3]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426070; PubMed=12537573;
RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J.W., Svirskaas R.,
RA Patel S., Friese E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Ceiniker S.E.;
RT "The transposable elements of the Drosophila melanogaster euchromatin:
a genomic perspective.";
RL Genome Biol. 3:RESEARCH0084.1-RESEARCH0084.20(2002).
[4]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,

RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettecourt B.R., Ceiniker S.E., de Grey A.D.N.J., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
systematic review.";
RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
[5]
RP NUCLEOTIDE SEQUENCE.
RX Berkeley Drosophila Genome Project;
RA Ceiniker S., Carlson J., Wan K., Pfeiffer B., Friese E., George R.,
RA Hoskins R., Stapleton M., Pacle J., Park S., Svirskaas R., Smith E.,
RA Yu C., Rubin G.;
RT "Drosophila melanogaster release 4 sequence.";
RN Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
[6]
RP NUCLEOTIDE SEQUENCE.
RX FlyBase;
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
RL EMBL: AE003749; AAF56356.2; -; Genomic_DNA.
DR MEROPS: S09.016; -; Drosophila melanogaster.
DR Ensembl: CG3744; -; Drosophila melanogaster.
DR FlyBase: FBgn0039240; CG3744.
DR GO: GO:0003824; F: catalytic activity; IEA.
DR GO: GO:0008236; F: serine-type peptidase activity; IEA.
DR GO: GO:0006508; P: proteolysis and peptidolysis; IEA.
DR GO: GO:0004469; P: peptidase_S9.
DR InterPro: IPR001325; Peptidase_S9.
DR InterPro: IPR004379; Ser_estrfs.
DR Pfam: PF00930; DPPIV_N; 1.
DR Pfam: PF00326; Peptidase_S9; 1.
SQ SEQUENCE 1113 AA; 123881 MW; 20857E3B212DF2E4 CRC64;
Alignment Scores:
Pred. No.: 0.00235 Length: 1113
Score: 13.00 Matches: 13
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 1.3% Indels: 0
DB: 2 Gaps: 0
US-10-825-632-2 (1-3120) x Q9VC19_DROME (1-1113)
QY 2416 ATCCACGGCTGGTCCTATGGAGGATACCTCTCCCTGATG 2454
Db 977 ILeHisGlyTrpSerTyrGlyGlyTyrLeuSerLeuMet 989
RESULT 14
O75273 HUMAN
ID O75273 HUMAN PRELIMINARY; PRT; 508 AA.
AC O75273
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE R26984.1 (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Lamerdin J.E., McCready P.M., Skowronski E., Adamson A.W.,
RA Burkhardt-Schultz K., Gordon L., Kyle A., Ramirez M., Stiliwagen S.,
RA Phan H., Velasco N., Do L., Regala W., Terry A., Ganes J., Avila J.,
RA Danganan L., Poundstone P., Christensen M., Georgescu A., Avila J.,
RA Liu S., Attix C., Andreise T., Trankheim M., Amico-Keller G.,
RA Cosfield J., Duarte S., Lucas S., Bruce R., Thomas P., Quan G.,
RA Kronmiller B., Arellano A.S., Montgomery M., Ow D., Nolan M., Trong S.,
RA Kobayashi A., Olsen A.S., Carraro A.V.;
RN Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
RL EMBL: AC005594; AAC33801.1; -; Genomic_DNA.

Mon May 8 13:41:51 2006

DR MEROPS; S09.019; -
DR Ensembl; ENSG0000142002; Homo sapiens.
DR GO; GO:0003824; F: catalytic activity; IEA.
DR InterPro; IPR001375; Peptidase_S9
DR InterPro; IPR002469; Peptidase_S9B.
DR InterPro; IPR000379; Ser_estrs.
DR Pfam; PF00930; DPPIV N; 1.
DR Pfam; PF00326; Peptidase_S9; 1.
FT NON TER 1 1
SQ SEQUENCE 508 AA; 57750 MW; 2F4D6645BE2D2C89 CRC64;

Alignment Scores:
Pred. No.: 0.313 Length: 508
Score: 11.00 Matches: 11
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 1.1% Indels: 0
DB: 2 Gaps: 0

US-10-825-632-2 (1-3120) x O75273_HUMAN (1-508)

QY 1804 CCITTAGACATCACCTGACGTAGTCAGTTAC 1836
Db 167 ProLeuGluHisIseuTyValValSerTyr 177

RESULT 15
Q6H9E3_9TRYP
ID Q6H9E3_9TRYP PRELIMINARY; PRT; 850 AA.
AC Q6H9E3;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Trypanosoma brucei dipeptidyl peptidase.
GN Name=dpptb;
OS Trypanosoma brucei.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
OX NCBI_TaxID=5691;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Bastos I.M.D., Girard D., Santana J., Grellier P.;
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ617830; CA92394.1; -; mRNA.
DR GO; GO:0003824; F: catalytic activity; IEA.
DR GO; GO:0008236; F: serine-type peptidase activity; IEA.
DR GO; GO:0006508; P: proteolysis and peptidolysis; IEA.
DR InterPro; IPR001375; Peptidase_S9.
DR InterPro; IPR002469; Peptidase_S9B.
DR InterPro; IPR000379; Ser_estrs.
DR InterPro; IPR02035; VWF_A.
DR Pfam; PF00930; DPPIV N; 1.
DR Pfam; PF00326; Peptidase_S9; 1.
DR PRINTS; PR00453; VWFADOMAIN.
SQ SEQUENCE 850 AA; 93842 MW; FB1D793DD9939BD4 CRC64;

Alignment Scores:
Pred. No.: 0.299 Length: 850
Score: 11.00 Matches: 11
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 1.1% Indels: 0
DB: 2 Gaps: 0

US-10-825-632-2 (1-3120) x Q6H9E3_9TRYP (1-850)

QY 2524 TATGATACAGGATACACGGAACGTTATATCGGT 2556
Db 731 TyrAspThrGlyTyrThrGluArgTyrMetGly 741

Search completed: May 2, 2006, 03:48:44
Job time : 725.5 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: May 2, 2006, 03:48:59 ; Search time 18.5 Seconds
(without alignments)
2788.628 Million cell updates/sec

Title: US-10-825-632-2
Perfect score: 1030
Sequence: 1 aagtgttaagctccgagg.....agaattactaaaaaaa 3120

Scoring table:

Xgapop 60.0, Xgapext 60.0
Ygapop 60.0, Ygapext 60.0
Fgapop 60.0, Fgapext 7.0
Delop 60.0, Delext 60.0

Searched: 572060 seqs, 82675679 residues

Word size:

Total number of hits satisfying chosen parameters: 1006954

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODEL=frame+n2p.model -DEV=xlp
-Q=/abss/ABSSWEB_Spool/US10825632/runat_01052006_110240_4511/app_query.fasta_1
-DB=Issued Patents AA -QPM=fastan -SUFFIX=oligo.n2p.ra -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=quality -THR MIN=1
-ALIGN=15 -MODE=LOCAL -OUTFWT=pto -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=2000000000 -HOST=abss02p
-USER=US10825632@cgn1_1_71 @runat_01052006_110240_4511 -NCPU=6 -ICPU=3
-NO MWAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=60 -FGAPEXT=7
-YGAPOP=60 -YGAPEXT=60 -DELOP=60 -DELEXT=60

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/1/iaa/5 COMB.pap.*
2: /cgn2_6/ptodata/1/iaa/6 COMB.pap.*
3: /cgn2_6/ptodata/1/iaa/H COMB.pap.*
4: /cgn2_6/ptodata/1/iaa/FCUTUS COMB.pap.*
5: /cgn2_6/ptodata/1/iaa/RE COMB.pap.*
6: /cgn2_6/ptodata/1/iaa/baCkfiles1.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	882	85.6	882	2	US-09-976-674-1
2	882	85.6	882	2	US-10-070-464-1
3	655	63.6	658	2	US-09-976-674-19
4	655	63.6	661	2	US-09-976-674-11
5	655	63.6	690	2	US-09-976-674-7
6	607	58.9	613	2	US-09-976-674-21
7	341	33.1	465	2	US-10-070-464-5
8	339	32.9	358	2	US-09-976-674-13
9	262	25.4	360	2	US-10-070-464-7
10	238	23.1	241	2	US-09-976-674-9
11	183	17.8	194	2	US-09-976-674-17

12	176	17.1	310	2	US-09-794-236-4	Sequence 4, Appli
13	176	17.1	310	2	US-10-070-464-3	Sequence 3, Appli
14	86	8.3	108	2	US-09-976-674-15	Sequence 15, Appl
15	41	4.0	54	2	US-09-621-976-5647	Sequence 5647, Ap
16	19	1.8	19	2	US-09-976-674-59	Sequence 59, Appl
17	13	1.3	159	2	US-09-270-767-32313	Sequence 32313, A
18	13	1.3	158	2	US-09-976-674-25	Sequence 25, Appl
19	13	1.3	819	2	US-09-976-674-37	Sequence 37, Appl
20	13	1.3	819	2	US-09-976-674-39	Sequence 39, Appl
21	13	1.3	832	2	US-09-976-674-29	Sequence 29, Appl
22	13	1.3	832	2	US-09-976-674-31	Sequence 31, Appl
23	13	1.3	863	2	US-09-976-674-33	Sequence 33, Appl
24	13	1.3	879	2	US-09-976-674-35	Sequence 35, Appl
25	13	1.3	892	2	US-09-976-674-23	Sequence 23, Appl
26	13	1.3	892	2	US-09-976-674-27	Sequence 27, Appl
27	9	0.9	30	2	US-09-205-258-1093	Sequence 1093, Ap
28	9	0.9	30	2	US-10-004-860-1093	Sequence 1093, Ap
29	9	0.9	97	2	US-09-205-258-1092	Sequence 1092, Ap
30	9	0.9	97	2	US-10-004-860-1092	Sequence 1092, Ap
31	9	0.9	133	2	US-09-716-129-84	Sequence 84, Appl
32	9	0.9	133	2	US-09-270-767-59553	Sequence 59553, A
33	9	0.9	222	2	US-09-270-767-44132	Sequence 44132, A
34	9	0.9	222	2	US-09-270-767-58740	Sequence 58740, A
35	9	0.9	315	2	US-09-716-129-62	Sequence 62, Appl
36	9	0.9	336	2	US-09-270-767-43390	Sequence 43390, A
37	9	0.9	564	2	US-09-248-796A-20544	Sequence 20544, A
38	9	0.9	564	2	US-09-196-270-2	Sequence 2, Appli
39	9	0.9	667	2	US-09-302-540-11280	Sequence 11280, A
40	9	0.9	1627	2	US-07-814-220-33	Sequence 33, Appl
41	8	0.8	25	1	US-07-814-220-33	Sequence 33, Appl
42	8	0.8	25	1	US-07-814-220-36	Sequence 36, Appl
43	8	0.8	28	1	US-07-812-421-36	Sequence 36, Appl
44	8	0.8	28	1	US-09-270-767-59966	Sequence 59966, A
45	8	0.8	28	2		

ALIGNMENTS

RESULT 1
US-09-976-674-1
; Sequence 1, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 882
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-1

Alignment Scores:
Pred. No.: 0
Score: 882.00
Percent Similarity: 100.0%
Best Local Similarity: 100.0%
Query Match: 85.6%
DB: 2

US-10-825-632-2 (1-3120) x US-09-976-674-1 (1-882)

QY 214 ATGGCAGCAGCAATGGAACAGAACAGCTGGGTGTGAGATATTTGAACACTCGGACTGT 273
|||||

Db 1 MetAlaAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
QY 274 GAGGAGAATATTGAATCACAGATCGCGCTAAATTCGAGCCCTTTTATGTTGAGCGGTAT 333
Db 21 GluGluAenIleGluSerGlnAspArgProGlyLeuGluProPheTyrrValGluArgTyr 40
QY 334 TCTCGAGCTCAGCTTAAAGAGCTGCTGCGATACAGAGAAATATATGCTACATGATG 393
Db 41 SerTrpSerGlnLeuLysLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
QY 394 CTAAGGACACACATGATTTTCATGTTTGTGAAGAGAAATGATCCAGATGGACCTCATCA 453
Db 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
QY 454 GACAGATCTATTACCTTGCATGCTCTGTGTGAGAACAGAGAAAAATACACTGTTTATCT 513
Db 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
QY 514 GAATTTCCAAAATCATCAATAGAGCAGCAGTCTTAATGCTCTCTTGAAGCCCTTTTG 573
Db 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120
QY 574 GATCTTTTTCAGCAACACTGGACTATGGAATGTATTTCTCGAGAGAAAGAACTATTAGA 633
Db 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGlnLeuLeuArg 140
QY 634 GAAAGAAAACGCAATGGAAACAGTCGGAATTTGCTTCTACGATTATCACCAAGGAGTGA 693
Db 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
QY 694 ACATTTCTGTTTCAAGCCGGTAGTGGAATTTATACGTAAAGATGGAGGCCACRAGGA 753
Db 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
QY 754 TTTACGCAACAACTTTAGGCCCAATCTAGTGGAACACTAGTCTCCCAACATACGATG 813
Db 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
QY 814 GATCCAAAATATGCCCCCTGATCCAGACTCGATGCTTTTATACATAGCAACGATATT 873
Db 201 AspProLysCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
QY 874 TGGATCTTAACATCTGACCAAGAGAGAAAGAGACTCACTTATGTGCACAATGAGCTA 933
Db 221 TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu 240
QY 934 GCCAACATGGAAGAAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAAGAAGAA 993
Db 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu 260
QY 994 TTTGATAGATATTCTCGCTATTGCTGTTGCTCCAAAGCTGAAACAACTCCCGAGTGGT 1053
Db 261 PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly 280
QY 1054 AAAATCTTAGAATTTCTATATGAAGAAATGATGAATCTGAGTGGGAAATTTATTCATGTT 1113
Db 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal 300
QY 1114 ACATCCCTATGTTGAAACAAAGGAGGCGAGATTCATTCCGTTATCTTAAACAGGTACA 1173
Db 301 ThrSerProMetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThr 320
QY 1174 GCAAACTCAAGTCACTTTTAAAGATGTCAGAAATATGATGATGCTGAGAGAGATC 1233
Db 321 AlaAsnProLysValThrPheLysMetSerGluLeuMetIleAspAlaGluGlyArgIle 340
QY 1234 ATAGATGTCATAGTAAGCACTAATTCACCTTTTGAGATCTTATTTGAGAGATTTGAA 1293
Db 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
QY 1294 TATATTGCCAGAGCTGGATGACTCTCGAGGAGAAATATGCTGGTCCATCTCTACATGAT 1353
Db 361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380

QY 1354 CGCTCCAGAGACTCCCTACAGATAGTGTGTATCTACCTGAAATATTATATCCAGTAGAA 1413
Db 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
QY 1414 GATGATGTTATGGAAAGCGAGAGACTCATGTAGTCTAGTCCCTGATTCTGTGTGACGCCACTA 1473
Db 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
QY 1474 ATTATCTATGAAGAAAACACAGACATCTGTGATAAATATCATGACATCTTTTCATCTTTT 1533
Db 421 IleIleTyrGluGluThrThrAspIleTyrPheAsnIleHisAspIlePheHisValPhe 440
QY 1534 CCCCAGAGTCACCAAGAGAGAAATGAGTTTATTTTTCCTCTGAAATGCAGAAACAGGTTTC 1593
Db 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
QY 1594 CGTCATTTATACAAAATTTACATCTATTTTAAAGGAAAGCAAAATATAAAGCATCCAGTGGT 1653
Db 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
QY 1654 GGGCTGCTCTCCCAAGTGATTTCAAGTGTCTCTATCAAGAGGAGATAGCAATTTACCAGT 1713
Db 481 GlyLeuProAlaProSerAspPheLysCysProLysGluGluIleAlaIleThrSer 500
QY 1714 GGTGAATGGGAAGTCTTGGCGGATGGATCTATATATCCAAGTTGATGAATCAGAGG 1773
Db 501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArg 520
QY 1774 CTGTATTTTGAAGGACCCAAAGACTCCCTTTTAGAGCATCACCTGTACGTAGTCAGT 1833
Db 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSer 540
QY 1834 TAGCTAAATCCTGGAGAGTGACAAAGCTGACTGACCGTGGCTACTCACATTTCTGTGTC 1893
Db 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
QY 1894 ATCAGTCAGACTGTGACTCTTTTATAAGTAAAGTATAGTAAACCAAGAGAAATCCACACTGT 1953
Db 561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
QY 1954 GTGTCCCTTTTACAAAGCTATCAAGTCTCGAAGATGACCCAACTTGCAAAACAAAGAAATTT 2013
Db 581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
QY 2014 TGGGCCACCATTTTGGATTTCAGCAGGTCTCTTCTCTGACTATATCTCCTCCAGAAATTTTC 2073
Db 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe 620
QY 2074 TCTTTTGAAGTACTACTGATTTTACATTTGATGGATGCTCTACAGCCCTCATGATCTA 2133
Db 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
QY 2134 CAGCCTGGAAAGAAATATCCTACTGCTGTTCATATATATGCTGCTCAGGTCCAGTTG 2193
Db 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu 660
QY 2194 GTCAATAATCCGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTTAGCTCTCTAGGT 2253
Db 661 ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly 680
QY 2254 TATGCGTTGTAGTGATGACAAACAGGGGATCTCTGTCCAGAGGGCTTAAATTTGAAGGC 2313
Db 681 TyrValValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly 700
QY 2314 GCCTTTAAATATAAATGGGTCAAATAGAAATTCAGCATCAGGTGGAAGCACTCCAATAT 2373
Db 701 AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr 720
QY 2374 CTAGCTTCTCGATATGATTTTCATTGACTTAGATTGCTGGGCAATCCACGGCTCGTCTAT 2433
Db 721 LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrSerTyr 740

```
QY 2434 GGAGGATACCTCCCTGATGCATTAAATGCAGAGGTCAGATATCTTCAGGGTGTCTATT 2493
Db |||||||
QY 741 GYGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheAsgValAlaIle 760
Db |||||||
QY 2494 GCTGGGGCCCCAGTCACCTCTGTGGATCTTATGATACAGGATACACGGAACTTATATG 2553
Db |||||||
QY 761 AlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluAsgTyrMet 780
Db |||||||
QY 2554 GGTACCCCTGCAGACAGAACTGAGGCTATTACTTAGGATCTGGCCATGCAAGCAGAA 2613
Db |||||||
QY 781 GlyHisProAspGlnAsnGluGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu 800
Db |||||||
QY 2614 AAGTTCCTCTCAACCAATCGTTTACTGCTCTTACATGCTTTTCTGGAGTGAATGTC 2673
Db |||||||
QY 801 LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820
Db |||||||
QY 2674 CATTTTGCACATACCATGATATATTACTGAGTCTTTTGTAGGGCTGGAAGCCCATATGAT 2733
Db |||||||
QY 821 HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840
Db |||||||
QY 2734 TTACAGATCTATCTCAGGAGACACAGCATTAAGATTCCTGATCGGGAACATATAT 2793
Db |||||||
QY 841 LeuGlnIleTyrProGlnArgHisSerIleArgValProGluSerGlyGluHisTyr 860
Db |||||||
QY 2794 GAATGTCATCTTTTGCACCTACCTCAAGAAACCTTGGATCACGATTTTGTCTCTAAAA 2853
Db |||||||
QY 861 GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaLeuLys 880
Db |||||||
QY 2854 GTGATA 2859
Db |||||||
QY 881 ValIle 882
Db |||||||

RESULT 2
US-10-070-464-1
; Sequence 1, Application US/10070464
; Patent No. 6881564
; GENERAL INFORMATION:
; APPLICANT: ABBOTT, Catherine Anne
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES
; FILE REFERENCE: GH-007
; CURRENT APPLICATION NUMBER: US/10/070,464
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: PCT/AU00/01085
; PRIOR FILING DATE: 2000-09-11
; PRIOR APPLICATION NUMBER: AU P05709
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: AU P02762
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 882
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-070-464-1

Alignment Scores:
Pred. No.: 0 Length: 882
Score: 882.00 Matches: 882
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 85.6% Indels: 0
DB: 2 Gaps: 0

US-10-825-632-2 (1-3120) x US-10-070-464-1 (1-882)
QY 214 ATGGCAGCAGCAATGGAAACAGAACAGCTGGGTGTTGAGATATTTGAAACTCGCGACTGT 273
Db |||||||
QY 1 MetAlaAlaIleMetGluThrGlnGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
Db |||||||
QY 274 GAGGAGAAATATTGAATCACAGGATCGGCCTTAAATTCGAGCCCTTTTATGTTGAGCGGTAT 333
Db |||||||
```

```
Db 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluAsgTyr 40
QY 334 TCCTGGAGTCAGCTTAAAAAGCTGCTGCGCATACACAGAAAAATATCATGCTCATCATGATG 393
Db |||||||
QY 41 SerTrpSerGlnLeuLysLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
Db |||||||
QY 394 GCTAAGCCACCATGATGTTTATGTTGTGAAGAGGAATGATCCAGATGACACCTCATTC 453
Db |||||||
QY 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
Db |||||||
QY 454 GACAGAAATCTATTACCTGCTGCCATGCTGCTGAGAACACAGAGAAAAATACACTGTTTATTCT 513
Db |||||||
QY 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
Db |||||||
QY 514 GAAATTTCCCAAACTATCAATAGACAGCAGCTTTAATGCTCTCTTGGAGCCCTCTTTTG 573
Db |||||||
QY 101 GluIleProLysThrIleAsnArgAlaValLeuMetLeuSerTrpLysProLeuLeu 120
Db |||||||
QY 574 GATCTTTTTCAGGCAACACTGGACTATGGAATGATTTCTCGAGAAGAAGAACTATTAA 633
Db |||||||
QY 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg 140
Db |||||||
QY 634 GAAAGAAAAGCATTGGAACAGTCGGAATTGCTTCTTACGATTATTCACCAAGGAAGTGA 693
Db |||||||
QY 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
Db |||||||
QY 694 ACATTTCTGTTTCAAGCCGCTAGTGGAAATTTATCACTGTAAGATGAGAGGCCCAACAG 753
Db |||||||
QY 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
Db |||||||
QY 754 TTTACGCAACACCTTTAAGGCCCAATCTAGTGGAACTAGTTGCTCCCAACATACGGATG 813
Db |||||||
QY 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
Db |||||||
QY 814 GATCCAAAATTTATGCCCGCTGATCCAGACTGCTGATTTGCTTTTATACATAGCAACGATAT 873
Db |||||||
QY 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
Db |||||||
QY 874 TGGATATCTAACATCGTAACACAGAGAAAGAGAGACTCACTTATGTGCACAATGAGCTA 933
Db |||||||
QY 221 TrpIleSerAsnIleValThrArgGluGluAsgArgLeuThrTyrValHisAsnGluLeu 240
Db |||||||
QY 934 GCCAATATGGAAGAAGATGCCAGATCAGCTGGAGTCGCTACCTTTGCTTCCACAGAGAA 993
Db |||||||
QY 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu 260
Db |||||||
QY 994 TTTTCATAGATATTTCTGCTATTGCTGCTGCTCCAAAGCTGAAACCACTCCACAGTGGTGG 1053
Db |||||||
QY 261 PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly 280
Db |||||||
QY 1054 AAAATTTCTTAGAATTTCTATATGAAGAAATGATGAAATCTGAGGTGGAAATTTATCATGTT 1113
Db |||||||
QY 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal 300
Db |||||||
QY 1114 ACATCCCTTATGTTGGAACAGAGGGGAGATTCATCTCCGTTATCTTAAACACAGTACA 1173
Db |||||||
QY 301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr 320
Db |||||||
QY 1174 GCAATATCTTAAAGTCACCTTTTAAAGATGTGAGAAATAATGATTTGATGCTGGAAGGAAGATC 1233
Db |||||||
QY 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
Db |||||||
QY 1234 ATAGATGTCATAGATAAGGAACCTAAATTCACCTTTTGGAGATCTCTATTGGAAGGAGTTGAA 1293
Db |||||||
QY 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
Db |||||||
QY 1294 TATATTGCCAGAGCTGATCGACTCTCTGAGGGAATATGCTTGGTCCATCTCTACTAGAT 1353
Db |||||||
QY 361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
Db |||||||
QY 1354 CGCTCCCAAGACTCCCTTACAGATAGTGTGATCTCACTCTGAATTTATTTATCCAGTAGAA 1413
Db |||||||
QY 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
Db |||||||
```

QY 1414 GATGATGTTATGAAGGACGAGACTCAATTGAGTCAGTCCCTGATTTCTGTGACGCCACTA 1473
Db 401 AspAspValMetGluArgGlnArgLeuIleGluSerValThrProLeu 420
QY 1474 ATTATCTATGAAGAAACAAACAGACATCTGGATAAATATCATGACATCTTTTCATGTTTTT 1533
Db 421 IleIleTyrGluGluThrThrAspIleTyrPheAsnIleHisAspIlePheHisValPhe 440
QY 1534 CCCCAAGTACAGGAAGGAAATTTAGTTTATTTTGGCTCTGAAATGCAAAACAGGTTTC 1593
Db 441 ProGlnSerHisGluGluIleGluPheIlePheAsnIleHisAspIlePheHisValPhe 460
QY 1594 CGTCATTTATACAAATATACATCTATTTTAAAGGAAGCAATATAAAGCATCCAGTGGT 1653
Db 461 ArgHisLeuTyrIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
QY 1654 GGCCTGCTCTCCAGTGATTTCAAGTGTCTCTATCAAGAGGAGATAGCAATTTACCAAGT 1713
Db 481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluIleAlaIleThrSer 500
QY 1714 GGTGAATGGGAAGTCTTGGCCGGCATGGATCTAATATCCAAAGTTGATGAAGTCAGAAGG 1773
Db 501 GlyGluTrpGluValLeuGluArgHisGlySerAsnIleGlnValAspGluValArgArg 520
QY 1774 CTGTATATTTTGAAGGCCACCAAGACTCCCTTTTAGAGCATCACCTGTACGTAGTCAGT 1833
Db 521 LeuValTyrPheGluGluThrLysAspSerProLeuGluHisLeuTyrValValSer 540
QY 1834 TACGTAATCTCGAGAGGTGACAGGCTGACTGACCGTGGCTACTCACATTTCTGTGTC 1893
Db 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
QY 1894 ATCAGTCAGCAGCTGTGACTCTTTTATAGTAAGTATAGTAACACAGAGAATCCACACTGT 1953
Db 561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
QY 1954 GTGTCCCTTTTACAAGCTATCAAGTCTCTGAAGTAGCCCACTCTCCAAACAAGGATTT 2013
Db 581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
QY 2014 TGGGCCACCATTTTGGATTACAGAGTCTCTCTGACTATCTCTCTGACTACTCTCCAGAAATTTTC 2073
Db 601 TipAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe 620
QY 2074 TCTTTTGAAGTACTGATTTACATTTGATGGATGCTCTACAAAGCCTCATGATCTA 2133
Db 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
QY 2134 CACCTTGAAGAAATATCCTACTGTCTGTCTATATATGATGCTCTCAGGTGCAAGTTG 2193
Db 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu 660
QY 2194 GTGAATAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCCCTAGCTCTTAGGT 2253
Db 661 ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly 680
QY 2254 TATGCTGTGTATGATACACAAAGGGATCCCTGTCCACCGGGCTTAAATTTGAAGC 2313
Db 681 TyrValValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly 700
QY 2314 GCCTTTAAATATAAATGGTCAATAGAAATTCACCATCAGGTGAAGGACTCCCAATAT 2373
Db 701 AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr 720
QY 2374 CTAGCTTCTCGATATGATTTCAATTGACTTAGATCGTGTGGCATCCACCGCTGGTCCCTAT 2433
Db 721 LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyr 740
QY 2434 CGAGGATACCTCTCCCTGATGGCATTTAATGACAGAGGTTCAGATATCTTCAGGCTGCTATT 2493
Db 741 GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760

QY 2494 GCTGGGCCCCAGTCACTCTGTGATCTTCTATGATACAGGATACACGGAACGTTATATG 2553
Db 761 AlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMet 780
QY 2554 GGTCACTCTGACCGAATGAAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGACAA 2613
Db 781 GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu 800
QY 2614 AAGTTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGTTTCTCGATGAGAATGTC 2673
Db 801 LysPheProSerGluProAsnArgLeuLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820
QY 2674 CATTTTCCACATACCAGTATATTACTAGTCTTTTGTAGTGGCTTGGAAAGCCATATGAT 2733
Db 821 HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840
QY 2734 TTACAGATCTATCTCTCAGAGAGACACAGATAAGAGTTCTCTGAATCGGAGAACATTAT 2793
Db 841 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860
QY 2794 GAACCTGATCTTTTGGCACTACCTTCAAGAAACCTTGGATCAGCTATTGCTGCTCTAAA 2853
Db 861 GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaLeuLys 880
QY 2854 GTGATA 2859
Db 881 ValIle 882
RESULT 3
US-09-976-674-19
; Sequence 19 Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19
; LENGTH: 658
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-19
Alignment Scores:
Pred. No.: 0 Length: 658
Score: 655.00 Matches: 655
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 63.6% Indels: 0
DB: 2 Gaps: 0
US-10-825-632-2 (1-3120) x US-09-976-674-19 (1-658)
QY 214 ATGGCAGCAGCAATCGAAACAGACAGCTGGCTTGTGAGATATTTGAAACTGCGGACTGT 273
Db 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
QY 274 GAGGAGATATTGATCATCAGATCGGCCTTAATTTGAGCGCTTTTATGTTGAGCGGTAT 333
Db 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
QY 334 TCTCTGAGTCAGCTTTAAAGACTCTCTGCGCATACCAAGAAATATCATGCTACATGATG 393
Db 41 SerTrpSerGlnLeuLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60

```
QY 394 GCTAAGGCACACATGATTTTCATGTTTGTGAAGAGAAATGATCCAGATGGACCTCATTCOA 453
Db 61 AlaiysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
QY 454 GACAGATCTATTACCTTGCATGCTGCTGAGACAGAGAAATACACTGTTTATCTTCT 513
Db 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
QY 514 GAAATTTCCAAAACATCAATAGACGACGAGTCTTAATGCTCTCTTTGAAAGCCCTCTTTTG 573
Db 101 GluileProLysThrIleAsnArgAlaAlaValLeuMetLeuSerIlePheProLeuLeu 120
QY 574 GATCTTTTTCAGGCAACACCTGACTATGGAATGTATTTCTCGAGAGAAAGAACTATTAAAGA 633
Db 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg 140
QY 634 GAAAGAAACCGCAATTGGAACAGTCGGAATTTGCTTCTTACGATTAATACCAAGGAAGTGA 693
Db 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
QY 694 ACATTTCTGTTTCAAGCCGCTAGTGGAAATTTATCATCGTAAAGATGGAGGCCACAAAGGA 753
Db 161 ThrLeuPhePheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
QY 754 TTTACGCAACAACTTTAAGCCCAATCTAGTGGAACTAGTTGTCCTCCAAACATACGGATG 813
Db 181 PheThrGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
QY 814 GATCCAAATTTATGCCCGCTGATCCAGACTGCGATTGCTTTTATACATAGCAACGATATT 873
Db 201 AspProLysLeuCysProAlaAspProAspTyrPheIleAlaPheIleHisSerAsnAspIle 220
QY 874 TGGATATCTAACATCGTAACACAGAGAAAGAGACTCACATTATGTCCACAAATGAGCTA 933
Db 221 TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu 240
QY 934 GCCAACATGGAAGAAGATGCCAGATCAGCTGGAGTGCCTTACCTTTGTTCTCCAAAGAGAA 993
Db 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu 260
QY 994 TTTGATAGATATTCTGCTATTTGCTGATGCTGAGTGCCTTACCTTCCAAAGAGAGTGGT 1053
Db 261 PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly 280
QY 1054 AAAATTTCTAGAATTTCTATGAGAAATGATGATCTGAGCTGGAAATTTATTCATGTT 1113
Db 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal 300
QY 1114 ACATCCCTATGTTGGAACACAGAGGCGAGATTCATTCCTGTTATCTTAAACAGGTACA 1173
Db 301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr 320
QY 1174 GCAATCTCTAAGTCACCTTTTAAGATGTCAGAAATATGATTTGATCTTCTTGAAGAGTGGAA 1233
Db 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
QY 1234 ATAGATGTCATAGATAAGAACTAATTCACACCTTTTGAATTTCTATTGAAAGAGTGGAA 1293
Db 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
QY 1294 TATATTGCCAGCTGGATGGACTCTCGAGGGAANAATATGCTTGTGTCCTACTAGAT 1353
Db 361 TyrIleAlaArgAlaGlyTyrThrProGluGlyLysTyrAlaTyrPheIleLeuLeuAsp 380
QY 1354 CDTCCCGAGCTCCCTACAGATAGTGTGATCTCACCTGGAATTTATTTATCCAGTAGAA 1413
Db 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
QY 1414 GATGATCTTATGGAAGGACAGACTCATTGAGTCAGTCCTGATTCCTGTCGACGCCACTA 1473
Db 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
QY 1474 ATTATCTATGAAGAAACACACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTT 1533
```

```
Db 421 IleIleTyrGluGluThrThrAspIleThrPheIleAsnIleHisAspIlePheHisValPhe 440
QY 1534 CCCCAAGTCCACAAAGAGGAAATTTGATTTTATTTTGTCTCTGAATGCAAAACAGGTTTC 1593
Db 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
QY 1594 CGTCATTTATACAAAATTTACATCTATTTTAAAGGAAGCAATATATAACGATCCAGTGGT 1653
Db 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
QY 1654 GGGCTGCTGCTCCAAAGTGAATTTCAAGTGTCTCTATCAAGAGGAGATAGCAATTTACAGT 1713
Db 481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer 500
QY 1714 GGTGAATGGGAAGTCTTGTGCCCGCATGATCTAATATCAAGTTGATGATCAGAGG 1773
Db 501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArg 520
QY 1774 CTGTATATTTTGAAGGCAACAAAGACTCCCTTTTAGAGCATCACCTGTAGTGTAGT 1833
Db 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSer 540
QY 1834 TACGTAATCTCGAGAGGTGACAAAGCTGACCTGCGTGGCTACTCACATTTCTGTCTGC 1893
Db 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
QY 1894 ATCAGTCAGCACTGTGACTTTCTTTATAGTAAGTAGTAACCAAGAGAAATCCACTGT 1953
Db 561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
QY 1954 GTGTCCCTTTACAAAGTATCAAGTCTCTGAAGATGACCCAACTTGCAAAACAAAGAAATTT 2013
Db 581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
QY 2014 TGGGCCACCAATTTTGATTCAGAGTCTCTCTCTCTGACTATCTCTCCAGAAATTTTC 2073
Db 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe 620
QY 2074 TCTTTTGAAGTACTACTGGATTTACATGTATGGGATGCTCTACAAGCCTCATGATCTA 2133
Db 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
QY 2134 CAGCTCGAAAGAAATATCTCTACTGTCTGTCTCATATATGTTGTGT 2178
Db 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGly 655

RESULT 4
US-09-976-674-11
; Sequence 11, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 661
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-11

Alignment Scores: 0 Length: 661
Pred. No.: 655 Matches: 655
Score: 655.00
```

Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	63.6%	Indels:	0
DB:	2	Gaps:	0

US-10-825-632-2 (1-3120) x US-09-976-674-11 (1-661)

QY	214	ATGCGAGAGCAATGGAAACAGAACAGCTGGGTGGTGGATATTTGAAACTGCCGACTGT	273
DB	1	MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys	20
QY	274	GAGGGAATATGGAATCACAGGATCGGCCTAAATTTGAGCCCTTTTATTTGTTGAGCGGTAT	333
DB	21	GluGluAenIleGluSerGlnAspArgProLyLeuGluProPheTyrValGluArgTyr	40
QY	334	TCCTGGAGTCAGCTTAAAAAGCTGCTTGGCGATACCAGAAAAATATCATGGCTACATGATG	393
DB	41	SerTrpSerGlnLeuLySLeuLeuAlaAspThrArgLySLeuGlyTyrHisGlyTyrMetMet	60
QY	394	GCTAAGGCACCATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGGACCTCATTTCA	453
DB	61	AlaLySAlaProHisAspPheMetPheValLySArgAsnAspProAspGlyProHisSer	80
QY	454	CACAGAATCTATTACCTTGCATGCTCTGGTGAGAACAGAGAAAAATACACTGTTTTATTCT	513
DB	81	AspArgIleTyrTyrLeuAlaMetSerGlyGluAenArgGluAenThrLeuPheTyrSer	100
QY	514	GAATTTCCCAAAATCATCAATAGACGACGACTTAATGCTCTCTTGGAGCCCTCTTTTG	573
DB	101	GluIleProLySThrIleAenAlaAlaValLeuMetLeuSerTrpLySProLeuLeu	120
QY	574	GATCTTTTTCAGCAACACTGGACTATGGAATGTTATCTCGAGAGAGAACTATTAAAGA	633
DB	121	AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuLeuArg	140
QY	634	GAAGAAGAACCATGGAACAGTCGGAATGCTCTTACGATATATCAACAAAGGAAGTGA	693
DB	141	GluArgLySArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly	160
QY	694	ACATTCTGTTTCAAGCCGTAGTGGAAATTTATACGTAAAGATGGAGGCGCCACAAGGA	753
DB	161	ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLySAspGlyGlyProGlnGly	180
QY	754	TTTACGCAACAACTTTAAGGCCCAATCTAGTGGAAACTAGTTGTCCCAACATACGGATG	813
DB	181	PheThrGlnGlnProLeuArgProAenLeuValGluThrSerCysProAenIleArgMet	200
QY	814	GATCCAAAATTTATGCCCGCTGATCCAGACTGGATTGCTTTTATACATAGCAACGATATT	873
DB	201	AspProLySLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle	220
QY	874	TGGATATCTAACATCGTAACAGAGAAAGAGACTCACTTATGTGCACAATGAGCTA	933
DB	221	TrpIleSerAenIleValThrArgGluGluArgGLeuThrTyrValHisAenGluLeu	240
QY	934	GCCAAATGGAAGAAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAAGAGAA	993
DB	241	AlaAenMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu	260
QY	994	TTTCATAGATATTCTGGCTATTGGTGTCCAAAAGCTGAAACAACTCCCAAGTGGTGT	1053
DB	261	PheAspArgTyrSerGlyTyrTrpCysProLySAlaGluThrThrProSerGlyGly	280
QY	1054	AAAATCTTGAATTTCTATAGCAAGAAATGATGATCTGAGGTGGAAATTTATCATGTT	1113
DB	281	LySLeuLeuArgIleLeuTyrGluGluAenAspGluSerGluValGluIleIleHisVal	300
QY	1114	ACATCCCTATGTTGAAACAGGAGGCGAGATTCATTCCTCGTTATCTCTAAACAGGTACA	1173
DB	301	ThrSerProMetLeuGluThrArgAlaAspSerPheArgTyrProLySThrGlyThr	320
QY	1174	GCAATCTCTAAAGTCACTTTTAAAGATGTCAGAAAAATGATGATGCTGGAAGGATC	1233

DB	321	AlaAenProLySValThrPheLySMetSerGluIleMetIleAspAlaGluGlyArgIle	340
QY	1234	ATAGATGTCATAGATAAGGAACCTAAATCAACCTTTTGAGATTCTATTGAAAGGAGTTGAA	1293
DB	341	IleAspValIleAspLySLeuIleGlnProPheGluIleLeuPheGluGlyValGlu	360
QY	1294	TATATTCCAGAGCTGGATGGACTCTCTGAGGGAATAATATGCTTGGTCCATCTACTAGAT	1353
DB	361	TyrIleAlaArgAlaGlyTrpThrProGluGlyLySLeuAlaTrpSerIleLeuLeuAsp	380
QY	1354	CGCTCCAGAGCTCGCCTACAGATAGTGTGATCTCACCTGAAATTTATTTATCCCAAGTAGAA	1413
DB	381	ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu	400
QY	1414	GATGATGTTATGGAAGGCGAGAGACTCATTTGAGTCAGTCCCTGATTTCTGTGACGCCACTA	1473
DB	401	AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu	420
QY	1474	ATTATCTATCAAGAAACAAACAGACATCTCGATAAAATATCCATGACATCTTTTCATGTTT	1533
DB	421	IleIleTyrGluGluThrThrAspIleTrpIleAenIleHisAspIlePheHisValPhe	440
QY	1534	CCCCAAAGTCACGAAGAGGAATTTGAGTTTATTTTGGCTCTGATGCAAAACAGGTTTC	1593
DB	441	ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLySThrGlyPhe	460
QY	1594	CGTCATTTATACAAAATACATCTATTTTAAAGGAAGCAATATAACGATCCAGTGGT	1653
DB	461	ArgHisLeuTyrLySLeuThrSerIleLeuLySLeuSerLySLeuTyrLySArgSerSerGly	480
QY	1654	GGCTGCTCTCTCAAGTGAATTTCAAGTGTCTCTATCAAGAGGAGATAGCAATTTACAGT	1713
DB	481	GlyLeuProAlaProSerAspPheLySLeuProIleLySLeuGluIleAlaIleThrSer	500
QY	1714	GGTGAATGGAAAGTCTTGGCCGGCATGGATCTAATATCCAAAGTTGATGAAGTCAGAAG	1773
DB	501	GlyGluTrpGluValLeuGlyArgHisGlySerAenIleGlnValAspGluValArgArg	520
QY	1774	CTGTATATTTTCAAGGCGCCAAAGACTCCCTTTTAGAGCATCACCTGTACGTAGTCAGT	1833
DB	521	LeuValTyrPheGluGlyThrLySAspSerProLeuGluHisLeuTyrValValSer	540
QY	1834	TACATAATCTCGAGAGGTGACAGGCTGACTGACCGTGGCTACTCACATTTCTTGCTGC	1893
DB	541	TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys	560
QY	1894	ATCAGTCAGCACTGTGACTTCTTTTATAAGTAAAGTATAGTAAACCAAGAAATCCACACTGT	1953
DB	561	IleSerGlnHisCysAspPhePheIleSerLySLeuSerAsnGlnLySAsnProHisCys	580
QY	1954	GTGTCCCTTTTACAAGCTATCAAGTCTCTGAAGATGACCCAACTTGCAAAACAAAGCAATTT	2013
DB	581	ValSerLeuTyrLySLeuSerSerProGluAspAspProThrCysLySThrLySLeuPhe	600
QY	2014	TGGGCCCACTTTTGGATTTCAGCAGTCTCTCTCTGACTATCTCTCTCCAGAAATTTTC	2073
DB	601	TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIlePhe	620
QY	2074	TCCTTTGAAAGTACTCGATTTCATGTTGATGATGATGATGATGATGATGATGATGATGAT	2133
DB	621	SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLySLeuProHisAspLeu	640
QY	2134	CAGCTCGAAAGAAATATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2178
DB	641	GlnProGlyLySLeuTyrProThrValLeuPheIleTyrGlyGly	655

RESULT 5

US-09-976-674-7
; Sequence 7, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akineanya, Karen


```
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 690
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-976-674-7

Alignment Scores:
Pred. No.: 0
Score: 655.00
Length: 690
Percent Similarity: 100.0%
Matches: 655
Best Local Similarity: 100.0%
Conservative: 0
Query Match: 63.6%
Mismatches: 0
Indels: 0
Gaps: 0
DB: 2

US-10-825-632-2 (1-3120) x US-09-976-674-7 (1-690)

QY 214 ATGGCAGCAGCAATCGAAACAGACAGCTGGGTGTGAGATATTTGAAACTCGCGACTGT 273
Db 1 MetAlaAlaAlaMetGluThrGluGluLeuGluValGluLeuPheGluThrAlaAspCys 20
QY 274 GAGGAGAAATTTGAATCACAGGATCGGCCTAAATTTGAGCCCTTTTATGTTGAGCGGTAT 333
Db 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
QY 334 TCTGGAGTCAGCTTAAAGCTGCTTGGCCGATACAGAAATATATCATGGCTACATGATG 393
Db 41 SerTrpSerGlnLeuLysLysLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
QY 394 GCTAAGGACCACTATGATTTTCATGTTTGTGAAGAGAAATGATCCAGATGGACCTCATCA 453
Db 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
QY 454 GNCAGAACTTATTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 513
Db 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
QY 514 GAAATTCCTCAAACTATCAATAGACAGCAGCTCTTAATGCTCTCTTGGAGCCTCTTTTG 573
Db 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120
QY 574 GATCTTTTTCAGGCAACACTGACTGATGGAATGATTTCTCGAAGAAAGAACTATTAGA 633
Db 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg 140
QY 634 GAAAGAAACCGATTTGGAACAGTCGGAATTCCTTTCATGATATCACCAGGAAGTGGGA 693
Db 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
QY 694 ACATTTCTGTTTCAAGCCGCTAGTGGAAATTTATCATCGTAAAGATGGAGGGCCACAAGA 753
Db 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
QY 754 TTTAGCCAAACACTTTAAGCCCAATCTAGTGGAACTAGTTGTCCTCCCAACATCGGATG 813
Db 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
QY 814 GATCCAAATATATGCTCCGCTGATCCAGACTGGATGCTTTTATACATAGCAACGATATT 873
Db 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
QY 874 TGGATATCTTAACATCGTAACAGAGAAAGAGAGACTCATTTATGTGCACAAATGAGCTA 933
Db 221 TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisGlnGluLeu 240
```

```
QY 934 GCCAACATGAAGAAGATGCCAGATCAGTGGAGTCGCTACCTTTGTTCTTCCAAAGAGAA 993
Db 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu 260
QY 994 TTTGATAGATATTCTGGCTATTGGTGTCTCCAAAGCTGAAACAACTCCAGAGTGGTGT 1053
Db 261 PheAspArgTyrSerGlyTyrTrpTyrCysProLysAlaGluThrThrProSerGlyGly 280
QY 1054 AAAATTCTTAGAATTTCTATATGAAGAAATGATGATCTGAGGTGGAATTTATTCATGTT 1113
Db 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleHisVal 300
QY 1114 ACATCCCTCTATGTTGGAACAACAGAGCGCAGATTCATTCCGTTATCTCTAAACACAGTACA 1173
Db 301 ThrSerProMetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThr 320
QY 1174 GCAAACTCTAAAGTCACCTTTTAAAGATGTCAGAAATAATGATTGATGCTGGAAGAGGATC 1233
Db 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
QY 1234 ATAGATGTCATAGATAAGGAACCTAATCAACCTTTTGGAGATTTCTATTGTAAGGAGTTGAA 1293
Db 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
QY 1294 TATATTGCCAGAGCTGGATGGACTCTCTGAGGAAATATGCTTGTGCTCATCTACTAGAT 1353
Db 361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
QY 1354 CGCTCCAGACTCGCTTACAGATAGTGTGATCTCACCTGATTTATTTATTTATTTATTTATTT 1413
Db 381 ArgSerGlnThrArgLeuGluIleValLeuIleSerProGluLeuPheIleProValGlu 400
QY 1414 GATGATGTTATGGAAGAGCAGAGACTCATGTAGTCAGTCCCTGATCTGTGTGACGCCACTA 1473
Db 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
QY 1474 ATTATCTATGAAGAAACAACAGACATCTGATAAATATCATGACATCTTTTCATGTTTTT 1533
Db 421 IleIleTyrGluGluThrThrAspIleTyrIleAsnIleHisAspIlePheHisValPhe 440
QY 1534 CCCCAAAGTCAAGAAAGAGAAATTTGATTTATTTTTCCTCTGAATTCGAAACACAGTTTC 1593
Db 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
QY 1594 CGTCATTTATACAAAATATCATCTATTATTTAAAGAAAGCAAAATATAACGATCCAGTGGT 1653
Db 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
QY 1654 GGGCTGCTGCTCCCAAGTGTATTCNAGTGTCTCTATCAGAGAGAGATAGCAATTACCAAGT 1713
Db 481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer 500
QY 1714 GGTGAATGGGAAGTCTTTGGCCGCTATGATCTTAATATCCAAATGATGATGATGATGATG 1773
Db 501 GlyGluThrGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520
QY 1774 CTGGTATATTTTGAAGCCACCAAGACTCCCTTTTAGACATCATCTGTACCTAGTTCAGT 1833
Db 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisIleLeuTyrValValSer 540
QY 1834 TACGTAAATCTCGAGAGGTGACAAGCTGACTGACCGTGGCTACTACATTTCTGCTGCTGC 1893
Db 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
QY 1894 ATCAGTCAGACTGCTGTGACTCTCTTTATTAAGTAAGTATAGTAAACAGAGAGATCCACTGT 1953
Db 561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
QY 1954 GTGTCCTTTTACAGCTTATCAAGTCTCTGAGAGATGACCCCACTTGCAGAAACAAGGAATTT 2013
Db 581 ValSerLeuTyrLysLeuSerSerProGluAspProThrCysLysThrLysGluPhe 600
```

QY 2014 TGGGCCACCAATTTGGATTACAGAGTCCCTCTCTGACTACTACTCTCCACGAAATTTTC 2073
Db 601 TTPAlaThrIleLeuAaspSerAlaGlyProLeuProAaspTyrThrProGluIlePhe 620
QY 2074 TCTTTTGAAGTACTACTGGATTTACATTTGTATGGAGTCTCTACAGCCTCATGATCTA 2133
Db 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAaspLeu 640
QY 2134 CAGCTCGAAGAAATATCTTACTGTCTCTCTCATATATGGTGGT 2178
Db 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGly 655

RESULT 6

US-09-976-674-21
; Sequence 21, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 21
; LENGTH: 613
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-21

Alignment Scores:
Pred. No.: 0 Length: 613
Score: 607.00 Matches: 607
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 58.9% Indels: 0
DB: 2 Gaps: 0

US-10-825-632-2 (1-3120) x US-09-976-674-21 (1-613)

QY 214 ATGCAGCAGCAATGGAAACAGAACAGCTGGGTGTGAGATATTTGAAACTGCGGACTGT 273
Db 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAaspCys 20
QY 274 GAGCAGAAATTTGAATCACAGGATCGGCCTAAATTTGGAGCCTTTTATGTTGAGCGGTAT 333
Db 21 GluGluAenileGluSerGlnAaspArgProLysLeuGluProPheTyrValGluAsgTyr 40
QY 334 TCCTGGAGTCAGCTTAAAGCTGCTTCCGATACCAGAAATATCATGGCTACATGATG 393
Db 41 SerTrpSerGlnLeuLysLysLeuLeuAlaAaspThrArgLysTyrHisGlyTyrMetMet 60
QY 394 GCTAAGCACCACATGATTTCAATGTTGTGAGAGGAATGATCCAGATGGACCTCATCA 453
Db 61 AlaLysAlaProHisAaspPheMetPheValLysArgAenAaspProAaspGlyProHisSer 80
QY 454 GACAGATCTATTACTTGGCATGCTGTGTGAGAACAGAGAAATATACACTGTTTATCT 513
Db 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAenArgGluAenThrLeuPheTyrSer 100
QY 514 GAAATTTCCAAAATCTAATAGCAGCAGCTTAAATGCTCTCTTGGAGCCCTCTTTTG 573
Db 101 GluIleProLysThrIleAenArgAlaValLeuMetLeuSerTrpLysProLeuLeu 120
QY 574 GATCTTTTTCAGGCAACACTGGACTATGGAATGTATCTCGAGAAGAAGAACTATTAA 633
Db 121 AspLeuPheGlnAlaThrLeuAaspTyrGlyMetTyrSerArgGluGluLeuLeuArg 140

QY 634 GAAAGAAAACCATTTGGAAACAGTCGGAATTTGCTTCTTACGATTATCACCAGAAAGTGA 693
Db 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAaspTyrHisGlnGlySerGly 160
QY 694 ACATTTCTTGTTTCAAGCCGGTAGTGGAAATTTATCAGTAAAGATGGAGGGCCACAAGGA 753
Db 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAaspGlyGlyProGlnGly 180
QY 754 TTTTACGCAACCAACCTTTAAGGCCCAATCTAGTGGAAACTAGTTGTCCCAACACATCGGATG 813
Db 181 PheThrGlnGlnProLeuAsgProAenLeuValGluThrSerCysProAenileArgMet 200
QY 814 GATCCAAAATTTATGCCCGCTGATCCAGACTGAGTGTCTTTTATACATAGCAACCATATT 873
Db 201 AspProLysLeuCysProAlaAaspProAaspTrpIleAlaPheIleHisSerAenAaspIle 220
QY 874 TGGATATCTAACATCGTAACACAGAGAAAGAGGAGCTCACATTATGTGCACAAATGAGCTA 933
Db 221 TrpIleSerAenileValThrArgGluGluAsgLeuThrTyrValHisAenGluLeu 240
QY 934 GCCAACATGGAAGAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCTCAAGAGAA 993
Db 241 AlaAenMetGluGluAaspAlaAsgSerAlaGlyValAlaThrPheValLeuGlnGluGlu 260
QY 994 TTTGATAGATATTTCTGCTATTTGCTGTCTCCAAAGCTGAAACAACTCCAGCTGGTGGT 1053
Db 261 PheAaspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly 280
QY 1054 AAAATTTCTAGAAATTTCTATATGAAAGAAATGATGAAATCTGAGGTGGAAATTTTTCATGTT 1113
Db 281 LysIleLeuArgIleLeuTyrGluGluAenAaspGluSerGluValGluIleIleHisVal 300
QY 1114 ACATCCCTATGTTGGAACCAAGAGGGCAGATTCATTCCTGTTATCTCTAAACAGGTACA 1173
Db 301 ThrSerProMetLeuGluThrArgArgAlaAaspSerPheArgTyrProLysThrGlyThr 320
QY 1174 GCAATCTCTAAAGTCACTTTTAAAGATGTCAGAAATAATGATGATGCTGGAAGAGATC 1233
Db 321 AlaAenProLysValThrPheLysMetSerGluIleMetIleAepAlaGluGlyArgIle 340
QY 1234 ATAGATGTCATAGATAAGGAACCTAATTCACCTTTTGGAGATTTCTATTGGAAGGAGTTGAA 1293
Db 341 IleAaspValIleAaspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
QY 1294 TATATTCCAGACTCGATGGACTCTCTGAGGGAATAATGCTTGGTCCATCTACTAGAT 1353
Db 361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAasp 380
QY 1354 CGCTCCAGACTCGCCTACAGATAGTGTGATCTCACCTGAATTTATTTATCCAGTAGAA 1413
Db 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
QY 1414 GATGATGTTATGAAAGGAGAGACTCATGAGTCAGTGCCTGATTCGTGACGCCCTA 1473
Db 401 AspAaspValMetGluArgGlnArgLeuIleGluSerValProAaspSerValThrProLeu 420
QY 1474 ATTATCTATGAGAAACAAACAGACATCTGGATAATAATFCCATGACATCTTTTCATGTTTTT 1533
Db 421 IleIleTyrGluGluThrThrAspIleTrpIleAenileHisAaspIlePheHisValPhe 440
QY 1534 CCCCAAAGTCACGAAGAGAAATTCAGTTTATTTTTCCTCTGAATCAAAACAGGTTTC 1593
Db 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
QY 1594 CGTCATTTATACAAAATTAATCTATTTTAAAGGAAAGCAATAATAACCATCCAGTGGT 1653
Db 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
QY 1654 GGCTCCCTGCTCCCAAGTGTATTTCAAGTGTCTCTATCAAGAGAGAGATAGCAATTTACCAGT 1713
Db 481 GlyLeuProAlaProSerAaspPheLysCysProIleLysGluGluIleAlaIleThrSer 500
QY 1714 GGTGAATGGGAAGTTCTTGGCCGCATGGATCTAATATCCAAAGTTGATGAAGTCAGAAGG 1773

Db 501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValA-gArg 520
Qy 1774 CTGGTATATTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCCTGTACGTAGTCAGT 1833
Db 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSer 540
Qy 1834 TACGTAATCCTCGAGAGGTGACAGGCTGACTGACCGCTGCTACTCACATTCTTGTGTC 1893
Db 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
Qy 1894 ATCAGTCAGCACTGTGACTTCTTTATAGTAAGTATAGTAACACGAGAAATCCACACTGT 1953
Db 561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
Qy 1954 GTGTCCTTTACAGCTATCAAGTCTGAGAGTACCCCACTTCGCAAAACAAGGAATTT 2013
Db 581 ValSerLeuTyrLysLeuSerProGluAspAspProThrCysLysThrLysGluPhe 600
Qy 2014 TGGGCCACCACTTTGGATTCA 2034
Db 601 TrpAlaThrIleLeuAspSer 607

RESULT 7

US-10-070-464-5
; Sequence 5, Application US/10070464
; Patent No. 6881564
; GENERAL INFORMATION:
; APPLICANT: ABBOTT, Catherine Anne
; APPLICANT: GORRELL, Mark Douglas
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES
; FILE REFERENCE: GH-007
; CURRENT APPLICATION NUMBER: US/10/070,464
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: PCT/AU00/01085
; PRIOR FILING DATE: 2000-09-11
; PRIOR APPLICATION NUMBER: AU P05709
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: AU P02762
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 465
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-070-464-5

Alignment Scores:
Pred. No.: 1.04e-313 Length: 465
Score: 341.00 Matches: 341
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 33.1% Indels: 0
DB: 2 Gaps: 0

US-10-825-632-2 (1-3120) x US-10-070-464-5 (1-465)

Qy 1165 ACAGTACAGCAATCTTAAGTCATCTTTAAGATCTCAGAAATATGATTCATGCTGAA 1224
Db 1 ThrGlyThrAlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGlu 20
Qy 1225 GDAAGGATCATAGATGCATAGATAAGGAACCTAATCAACCTTTTGAGATCTTATTTGAA 1284
Db 21 GlyArgIleIleAspValIleAspLysGluLeuIleGlnProPheGluLeuPheGlu 40
Qy 1285 GGAGTTGAATATATTGCGACGCTGGATGACTCCTGAGGAAAAATATGCTTGGTCCATC 1344
Db 41 GlyValGluTyrIleAlaArgAlaGlyTyrThrProGluGlyLysTyrAlaTrpSerIle 60
Qy 1345 CTACTAGATCGCTCCAGACTCGGCTACAGATGTTGATCTCACCTGCAATTTATTC 1404
Db 61 LeuLeuAspArgSerGlnThrArgLeuGlnIleValIleLeuSerProGluLeuPheIle 80

Qy 1405 CCAGTAGAAGATCATGTTATGGAAGGACAGAGACTCATTTAGTGCAGTGCCTGATTCGTG 1464
Db 81 ProValGluAspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerVal 100
Qy 1465 ACCCCACTAATTTATCTATGAGAACACACACACATCTGGATATAAATATATCCATGACATCTTT 1524
Db 101 ThrProLeuIleIleTyrGluIleThrThrAspIleIlePheAsnIleHisAspIlePhe 120
Qy 1525 CATGTTTTTCCCCCAAGTCACGAGAGGAAATTTGAGTTTATTTTGGCTCTGAAATGAAA 1584
Db 121 HisValPheProGlnSerHisGluGluIleGluPheIlePheAlaSerGluCysLys 140
Qy 1585 ACAGGTTCCGTCATTTATACAAATTCATCTATTTTAAAGGAAAGCAATATAAACA 1644
Db 141 ThrGlyPheArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArg 160
Qy 1645 TCCAGTGGTGGGCTGCTCTCCAGTGTCTCAAGTGTCTTATCAAGAGAGGAGATAGCA 1704
Db 161 SerSerGlyGlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAla 180
Qy 1705 ATTACCAAGTGTGAATGGGAAGTTCTTGGCCGCAATGGATCTAATATCAAGTTGATGAA 1764
Db 181 IleThrSerGlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGlu 200
Qy 1765 GTCAGAAGGCTGTATATTTTGAAGGCAACAAAGACTCCCTTTAGAGCATCCTGTAC 1824
Db 201 ValArgArgLeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyr 220
Qy 1825 GTAGTCAGTTACGTAATCTCGAGAGGTGACAAAGCTGACTGACCGTGGCTACTTCACAT 1884
Db 221 ValValSerTyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHis 240
Qy 1885 TCTTGTGTCATCAGTCAGCACTGTGACTTCTTTATAGTAAGTATAGTAACCAAGAAAT 1944
Db 241 SerCysCysIleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsn 260
Qy 1945 CCACACTGTGTGCTTTCATCAAGCTATCAAGCTATCAAGTCTGAGATGACCCCACTGCAAAACA 2004
Db 261 ProHisCysValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThr 280
Qy 2005 AAGCAATTTGGCCACCACTTTTGGATTCAGCAGTCTCTTCTTCATGACTATCTCTCCA 2064
Db 281 LysGluPheTrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProPro 300
Qy 2065 GAAATTTTCTCTTTTGAAGTACTACTGGAATTTACATTTGATGGAATGCTCTACAAGCTT 2124
Db 301 GluIlePheSerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysPro 320
Qy 2125 CATGATCTACAGCTCGGAAAGAAATATCTTACTGTGCTGCTTCATATATGGTGGTCTCAG 2184
Db 321 HisAspLeuGlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyProGln 340
Qy 2185 GTG 2187
Db 341 Val 341

RESULT 8

US-09-976-674-13
; Sequence 13, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61

```
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-13

Alignment Scores:
Pred. No.: 8,51e-312 Length: 358
Score: 339.00 Matches: 339
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 32.9% Indels: 0
DB: 2 Gaps: 0

US-10-825-632-2 (1-3120) x US-09-976-674-13 (1-358)

QY 214 ATGGCAGCAGCAATGGAAAACAGACAGCTGGGTGTGTGAGATATTTGAAACTGCGACTGT 273
Db 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
QY 274 GAGGAGAAATTTGAATCACAGGATCGGCCTAAATTTGGAGCCCTTTTATGTTGACGGTAT 333
Db 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyValGluArgTyr 40
QY 334 TCCTGGAGTCAGCTTAAAAAGCTTGCTGCGATACCAAGAAATATCATGGCTACATCATG 393
Db 41 SerTrpSerGlnLeuLysLysLeuLeuAlaAspThrArgLysGlyHisGlyTyrMetMet 60
QY 394 GCTAAGGCACACATGATTTTCATGTTTGTGAAGAGAAATGATCCAGATGGACCTCATTC 453
Db 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
QY 454 GACAGAATCTATTACCTTGCCATGCTGCTGAGACAGACAGAAATACACTGTTTATCT 513
Db 81 AspArgIleTyrIleuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
QY 514 GAATATCCCAAACTATCAATAGACAGCAGCTTAAAGCTCTCTTGGAGCCCTCTTTTG 573
Db 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120
QY 574 GATCTTTTTCAGGCACACTGGACTATGGAATGTAATTCGAGAAGAAAGAACTATTAAGA 633
Db 121 AspleuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg 140
QY 634 GAAAGAAACGCAATTCGAACAGTCGGAATTCCTTACGATTATCAACCAAGGAAGTGA 693
Db 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
QY 694 ACATTTCTGTTTCAAGCCGGTAGTGAATTTTATCACTAAAGATGGAGGCCACCAAGGA 753
Db 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
QY 754 TTTACGCAACACCTTTTAAGGCCCAATCTAGTGAACACTAGTTGTCCTCCCAACATACGATG 813
Db 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
QY 814 GATCCAAAATTTATGCCCCCTGATCCAGACTGGATTGCTTTTATACATAGCAACGATATT 873
Db 201 AspProLysLeuLysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
QY 874 TGGATATCTAACATCTAACACAGAGAAGAGAGACTCACTTATGTCAGATGAGCTA 933
Db 221 TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisGlnLeu 240
QY 934 GCCAATCATGAGAGATGCCAGATCAGCTGAGCTCGCTACCTTTGTTCTCCAGAGAA 993
Db 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu 260
QY 994 TTTGATAGATATCTGGCTATGTTGGTGTCCAAAAGCTGAACCACTCCAGTGTGTGT 1053
Db 261 PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly 280
```

```
QY 1054 AAAATTCCTAGATTCCTATATGAAAGAAATGATGAATCTGAGGTGAAATATTATCATGTT 1113
Db 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleHisVal 300
QY 1114 ACATCCCTATGTTGGAACACAGAGGCGAGATTTCATTCGTTATCTCTAAACAGGTACA 1173
Db 301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr 320
QY 1174 GCAATCTCTAAAGTCACTTTTAAAGATGTCAGAAATAATGATTGATGCTGAAAGGAGG 1230
Db 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArg 339
RESULT 9
US-10-070-464-7
; Sequence 7, Application US/10070464
; Patent No. 6881564
; GENERAL INFORMATION:
; APPLICANT: ABBOTT, Catherine Anne
; APPLICANT: GORRELL, Mark Douglas
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES
; FILE REFERENCE: GH-007
; CURRENT APPLICATION NUMBER: US/10/070,464
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: PCT/AU00/01085
; PRIOR FILING DATE: 2000-09-11
; PRIOR APPLICATION NUMBER: AU PQ5709
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: AU PQ2762
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 360
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-070-464-7

Alignment Scores:
Pred. No.: 5.89e-239 Length: 360
Score: 262.00 Matches: 262
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 25.4% Indels: 0
DB: 2 Gaps: 0

US-10-825-632-2 (1-3120) x US-10-070-464-7 (1-360)

QY 1756 GTTGATGAAGTCAGAAAGGCTGGTATATTTTCAAGGCACCAAGACTCCCTTTAGAGCAT 1815
Db 99 ValAspGluValArgArgLeuValTyrPheGluGlyThrLysAspSerProLeuGluHis 118
QY 1816 CACCTGTAGTGTAGTTCAGTTACGTAATCTCGAGAGGTGACAGGCTGACTGACCGTGC 1975
Db 119 HisLeuTyrValValSerTyrValAsnProGlyValThrArgLeuThrAspArgGly 138
QY 1876 TACTCACATTCCTGCTGCATCAGTCAGCAGCTGCTCTTTATAGTAGTATAGTAAC 1935
Db 139 TyrSerHisSerCysCylIleSerGlnHisCysAspPheIleSerLysTyrSerAsn 158
QY 1936 CAGAGAATCCACACTGTGCTGCTCTTACAGCTATCAAGTATCCTGAGATGCCCACT 1995
Db 159 GlnLysAsnProHisCysValSerLeuTyrLysLeuTyrSerProGluAspAspProThr 178
QY 1996 TGCAAAACCAAGAAATTTTGGGCCACCATTTTGGATTCAGCAGCTCTCTCTCTGACTAT 2055
Db 179 CysLysThrLysGluPheTrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyr 198
QY 2056 ACTCTCCAGAAATTTTCTCTTTTGAAGTACTACTCGATTACATTTGATGGGATGCTC 2115
Db 199 ThrProProGluIlePheSerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeu 218
QY 2116 TACAAGCCTCATGATCTACAGCCTGGAAGAAATATCTTACTGTGCTTCTATATATGT 2175
```

Db 219 TyrIysProHisAspLeuGlnProGlyLysLysTyrProThrValLeuPheIleTyrGly 238
QY 2176 GGTCTCTCAGGTGCGAGTAAATCGGTTTAAAGGAGTCAAGTATTTCCGCTGAAT 2235
Db 239 GlyProGlnValGlnLeuValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsn 258
QY 2236 ACCCTAGCTCTCTAGGTTATGTTGTTAGTATAGACACAGGGGATCTGTCCACGA 2295
Db 259 ThrLeuAlaSerLeuGlyTyrValValValIleAspAsnArgGlySerCysHisArg 278
QY 2296 GGGCTTAAATTTGAAGCGCTTTAAATATAAATGAGTCAAAATGAGCATCAG 2355
Db 279 GlyLeuLysPheGluGlyAlaPheLysTyrLysMetGlyGlnIleGluLeuAspGln 298
QY 2356 GTGAAGAGCTCCAATATCTAGCTTCTCGATATGATTTCAATGACTTAGATCGTGGC 2415
Db 299 ValGluGlyLeuGlnTyrLeuAlaSerArgTyrAspPheIleAspLeuAspArgValGly 318
QY 2416 ATCCACGGCTGGTCTATGAGGATACCTCTCCCTGATGGCATTAATCCAGAGTCCAGAT 2475
Db 319 IleHisGlyTrpSerTyrGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAsp 338
QY 2476 ATCTTCAGGGTGTCTATTGCTGGGCCCCAGTCACTCTCTGATCTCTTATGATACAGGA 2535
Db 339 IlePheArgValAlaIleAlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGly 358
QY 2536 TACACG 2541
Db 359 TyrThr 360
RESULT 10
US-09-976-674-9
; Sequence 9, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976, 674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 241
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-9
Alignment Scores:
Pred. No.: 3.18e-216 Length: 241
Score: 238.00 Matches: 238
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 23.1% Indels: 0
DB: 2 Gaps: 0
US-10-825-632-2 (1-3120) x US-09-976-674-9 (1-241)
QY 214 ATGCACGACCAATGGAACAGACAGCTGGGTGTTGAGATATTTGAAACTGCGACTGT 273
Db 1 MetAlaAlaLeuMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
QY 274 GAGGAGATATTTGAATCACAGGATCGGCCTTAATTTGGAGCCCTTTTATGTTGAGCGTAT 333
Db 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
QY 334 TCCTGGAGTACGCTTAAAGCTGCTTGGCGATACACAGAAATATCATGCTACATGATG 393

Db 41 SerTrpSerGlnLeuLysLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
QY 394 GCTAAGGCACACATGATTTTCATGTTTGTCAAGAGAAATGATCCAGATGGACCTCATCA 453
Db 61 AlalysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
QY 454 GACAGATCTATTACCTTGCCATGCTGCTGAGAACAGAGAAATACACTGTTTATTTCT 513
Db 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
QY 514 GAAATTTCCAAAACATCATCAATAGACGAGCTTAAATGCTCTCTTTGGAAGCTCTTTTG 573
Db 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120
QY 574 GATCTTTTTCAGCAACACATGGAATGGAATGATTTATTCGAGAGAGAACTATTAAGA 633
Db 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg 140
QY 634 GAAAGAAAACGATTCGACAGTCGAAATGCTTCTTACGATTATCATCAGGAGGAGTGA 693
Db 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
QY 694 ACATTTCTGTTTCAAGCCGCTAGTGAATTTATCAGTAAAGATCGAGGCGCAACAGGA 753
Db 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
QY 754 TTTACGCAACACCTTTAAGCCCAATCTAGTGGAAACTAGTTGTCCTCCCAACATACGATG 813
Db 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
QY 814 GATCAAAATATGCCCCGCTGATCCAGACTGGATGCTTTTATACATAGCAACGATATT 873
Db 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
QY 874 TGGATATCTAACATCGTAACAGAGAGAAAGAGAGACTCACATTATGTGCACAAAT 927
Db 221 TrpIleSerAsnIleValThrArgGluGluArgArgLeuThrTyrValHisAsn 238
RESULT 11
US-09-976-674-17
; Sequence 17, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976, 674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17
; LENGTH: 194
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-17
Alignment Scores:
Pred. No.: 3.52e-164 Length: 194
Score: 183.00 Matches: 183
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 17.8% Indels: 0
DB: 2 Gaps: 0
US-10-825-632-2 (1-3120) x US-09-976-674-17 (1-194)
QY 214 ATGCACGACCAATGGAACAGACAGCTGGGTGTTGAGATATTTGAAACTGCGACTGT 273

Db 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
QY 274 GAGGAGAAATATTGAATCACAGGATCGGCCTAAATTCGAGCCCTTTTATGTGTGAGCGGTAT 333
Db 21 GluGluAenIleGluSerGlnAspArgProLysLeuGluProPheTy-ValGluArgTyr 40
QY 334 TCCTGGAGTCAGCTTAAAAAGCTGCTTGGCGATACCAGAAAAATATCATGGCTACATGATG 393
Db 41 SerTrpSerGlnLeuLysLysLeuLeuAlaAspThrArgLysTy-HisGlyTyMetMet 60
QY 394 GCTAAGGACACCATGATTTTCATGTTTGTGAGAGAAATGATCCAGATGGACCTCATTC 453
Db 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
QY 454 GACAGATCTATTACCTTGCATGCTGCTGGAGACAGAGAAATACACTGTTTATCT 513
Db 81 AspArgIleTyTyLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTySer 100
QY 514 GAATTTCCCAAACTATCAATAGACGACGAGTCTTAATGCTCTCTTGAAGCCCTTTTG 573
Db 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120
QY 574 GATCTTTTTCAGCAACACTGGACTATGGAATGATTTCTCGAGAGAAAGAACTATTAA 633
Db 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTySerArgGluGluLeuLeuArg 140
QY 634 GAAAGAAAAGCATTGGAACAGTCGGAATTTGCTTACGATTATCAACCAAGGAAGTGA 693
Db 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyAspTyrHisGlnGlySerGly 160
QY 694 ACATTTCTGTTTCAAGCGGCTAGTGGAAATTTATACGTAAAGATGGAGGCGCCACAAGGA 753
Db 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyHisValLysAspGlyGlyProGlnGly 180
QY 754 TTTACGCCAA 762
Db 181 PheThrGln 183

RESULT 12

US-09-794-236-4
; Sequence 4, Application US/09794236
; Patent No. 6337069
; GENERAL INFORMATION:
; APPLICANT: Grouzmann, Eric
; APPLICANT: Lacroix, Jean-Silvain
; APPLICANT: Monod, Michel
; TITLE OF INVENTION: Method of Treating Rhinitis and Sinusitis
; FILE REFERENCE: 81985/276823
; CURRENT APPLICATION NUMBER: US/09/794,236
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 4
; LENGTH: 310
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-794-236-4

Alignment Scores:
Pred. No.: 1,37e-157 Length: 310
Score: 175.00 Matches: 176
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 17.1% Indels: 0
DB: 2 Gaps: 0

US-10-825-632-2 (1-3120) x US-09-794-236-4 (1-310)

QY 2332 GGTCAATAGAAATTCACGATCAGGTGGAGAGCTCCAAATATCTAGCTTCTCGATATGAT 2391
Db 135 GlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyLeuAlaSerArgTyAsp 154
QY 2392 TTCATTGACTTAGATCGTGTGGGCATCCACGGCTGGTCTTATGGAGGATACCTCTCCCTG 2451

Db 155 PheIleAspLeuAspArgValGlyIleHisGlyTrpSerTy-GlyTyLeuSerLeu 174
QY 2452 ATGCATTATTCAGAGGTCAGATATCTTCAGGTTGCTATTGCTGGGGCCCCCAGTCACT 2511
Db 175 MetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAlaGlyAlaProValThr 194
QY 2512 CTGTGGATCTCTATGATACAGGATACAGGACGTTATATGGTCACCTGACCCAGAT 2571
Db 195 LeuTrpIlePheTyAspThrGlyTyThrGluArgTyRwecGlyHisProAspGlnAsn 214
QY 2572 GAAAGGGCTATTACTTAGGATCTGTGGCATCGAAGCAGAAAAAGTTCCCTCTGGAACA 2631
Db 215 GluGlnGlyTyTyLeuGlySerValAlaMetGlnAlaGluLysPheProSerGluPro 234
QY 2632 AATCGTTACTGCTTTACATGTTCTCTGGATGAGAATGTCCATTTTGCACATACCACT 2691
Db 235 AsnArgLeuLeuLeuLeuHisGlyPheLeuAspGluAsnValHisPheAlaHisThrSer 254
QY 2692 ATATTACTAGTTTTTTAGTGAGGCTGGAAAGCCATATGATTTACAGATCTATCCTCAG 2751
Db 255 IleLeuLeuSerPheLeuValArgAlaGlyLysProTyAspLeuGlnIleTyProGln 274
QY 2752 GAGAGACACAGCATAAGAGTTCTCTGAATCGGAGAACATTATGAACCTGCATCTTTGAC 2811
Db 275 GluArgHisSerIleArgValProGluSerGlyGluHisTyGluLeuHisLeuLeuHis 294
QY 2812 TACCTTCAAGAAAACCTTGGATCCAGTATTGCTGCTCTAAAGTGATA 2859
Db 295 TyrLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeuLysValIle 310

RESULT 13

US-10-070-464-3
; Sequence 3, Application US/10070464
; Patent No. 6881564
; GENERAL INFORMATION:
; APPLICANT: ABBOTT, Catherine Anne
; APPLICANT: GORRELL, Mark Douglas
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES
; FILE REFERENCE: GH-007
; CURRENT APPLICATION NUMBER: US/10/070,464
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: PCT/AU00/01085
; PRIOR FILING DATE: 2000-09-11
; PRIOR APPLICATION NUMBER: AU PQ5709
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: AU PQ2762
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 310
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-070-464-3

Alignment Scores:
Pred. No.: 1,37e-157 Length: 310
Score: 175.00 Matches: 176
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 17.1% Indels: 0
DB: 2 Gaps: 0

US-10-825-632-2 (1-3120) x US-10-070-464-3 (1-310)

QY 2332 GGTCAATAGAAATTCACGATCAGGTGGAGAGCTCCAAATATCTAGCTTCTCGATATGAT 2391
Db 135 GlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyLeuAlaSerArgTyAsp 154
QY 2392 TTCATTGACTTAGATCGTGTGGGCATCCACGGCTGGTCTTATGGAGGATACCTCTCCCTG 2451
Db 155 PheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyGlyTyLeuSerLeu 174

```
QY 2452 ATGGCATTAAATGACAGAGTTCAGATATCTTACAGGTTCTCTATTCTCTGGGGCCCCAGTCACT 2511
Db 175 MetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAlaGlyAlaProValThr 194
QY 2512 CTGTGATCTTCTATGATACAGGATACACGGAAGTTATATGGTTCACCTGACACAGAAAT 2571
Db 195 LeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMetGlyHisProAspGlnAsn 214
QY 2572 GACAGGGCTATTACTTAGTCTGTGCCATGTCAAGCAGAGAAAGTTCCCTCTGAACCA 2631
Db 215 GluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGluLysPheProSerGluPro 234
QY 2632 AATCGTTTACTGCTCTTACATGTTTCTCGATGAGAAATGTCATTTTGCACATACCAAT 2691
Db 235 AsnArgLeuLeuLeuLeuHisGlyPheLeuAspGluAsnValHisPheAlaHisThrSer 254
QY 2692 ATATTACTAGTTTTTATGAGGCTGGAAGCCATATGATTTACAGATCTATCTCAG 2751
Db 255 IleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAspLeuGlnIleTyrProGln 274
QY 2752 GAGACACAGCATAGAGTCTCTGAATCGGAGAACATTTATGAATGCACTCTTTTGCAC 2811
Db 275 GluArgHisSerIleArgValProGluSerGlyGluHisTyrGluLeuHisLeuLeuHis 294
QY 2812 TACCTTCAAGAAACCTTCGATCAGCTATTGCTGCTCTTAAAGTGATA 2859
Db 295 TyrLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeuLysValIle 310
RESULT 14
US-09-976-674-15
; Sequence 15, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15
; LENGTH: 108
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-15
Alignment Scores:
Pred. No.: 2,23e-72 Length: 108
Score: 86.00 Matches: 86
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 8.3% Indels: 0
Gaps: 2
Db:
US-10-825-632-2 (1-3120) x US-09-976-674-15 (1-108)
QY 214 ATGGCAGCAGCAATGGAACAGAACACGTGGGTGTGAGATATTTGAATCTCGGACTGT 273
Db 1 MetAlaAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
QY 274 GAGGAGATATTGAATTCACAGATCGGCCTAAATTTGGAGCCTTTTATGTTGAGCGGTAT 333
Db 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
QY 334 TCCTGGAGTCAAGCTTAAAGCTCTTGGCCGATACAGAAATATCATGGCTACATGATG 393
Db 41 SerTrpSerGlnLeuLysLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
```

```
QY 394 GCTAAGGCACCATGATTTTCATGTTTGAAGAGGAATGATCCAGATGGACCTCATTTCA 453
Db 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
QY 454 GACAGAATCTATTACCTT 471
Db 81 AspArgIleTyrTyrLeu 86
RESULT 15
US-09-621-976-5647
; Sequence 5647, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 5647
; LENGTH: 54
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -39..-1
US-09-621-976-5647
Alignment Scores:
Pred. No.: 9,24e-30 Length: 54
Score: 41.00 Matches: 41
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 4.0% Indels: 0
Gaps: 2
Db:
US-10-825-632-2 (1-3120) x US-09-621-976-5647 (1-54)
QY 2602 ATGCAAGCAGAAAAAGTTCCCTCTGAACCAATCGTTTACTGCTCTTACATGTTTCTTG 2661
Db 1 MetGlnAlaGluLysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeu 20
QY 2662 GATGAGAATGTCATTTTGCACATACATACAGTATATTACTGAGTATTTTGTAGTCAGGCTGGA 2721
Db 21 AspGluAsnValHisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGly 40
QY 2722 AAG 2724
Db 41 Lys 41
Search completed: May 2, 2006, 03:55:26
Job time : 128.5 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame plus n2p model

Run on: May 2, 2006, 04:41:36 ; Search time 85.9 Seconds
(without alignments)

Title: US-10-825-632-2

Perfect score: 1030

Sequence: 1 aagtgcctaaagcctccgagg.....agaattactaaaaaaaaa 3120

Scoring table: OLICO

Xgapop	60.0	Xgapext	60.0
Ygapop	60.0	Ygapext	60.0
Fgapop	60.0	Fgapext	7.0
Delop	60.0	Delext	60.0

Searched: 1867569 seqs, 417829326 residues

Word size: 32 bit

Total number of hits satisfying chosen parameters: 3572506

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Command line parameters:

```
-MODEL=frame& n2p_model -DEV=xylp
-O/-O/b66/ABESSE/ABESSE_pool/US10825632/runat 01052006 110247 4670/app_query.fasta_1
-DB=Published Applications AA_Main -QFMT=fastan -SUFFIX=oligo.n2p.rapbm
-MINMATCH=0.1 -LOOCPUL=0 -LOOPEXT=0 -UNITSIZE=0 -START=1 -END=1 -MATRIX=oligo
-TRANS=human4.0.cdi -LIST=45 -DOCALLAG=200 -THR SCORE=quality -THR MIN=1
-ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -HOST=abs02p
-USER=US10825632 @CGN 1 1 307 @runat 01052006 110247 4670 -NCPU=6 -ICPU=3
-NO MAP -NEG SCORES -WAIT -DSPBLCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -FGAPEXT=60 -FGAPEXT=7
-XGAPOP=60 -YGAPEXT=60 -DELOP=60 -DFLEXT=60
```

```
Database : Published Applications_AA_Main:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SIMMARTES

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	882	85.6	882	3	US-09-976-674-1	Sequence 1, Appli
2	882	85.6	882	4	US-10-054-776-2	Sequence 2, Appli
3	882	85.6	882	4	US-10-070-789-38	Sequence 38, Appl
4	882	85.6	882	4	US-10-311-035-9	Sequence 9, Appli
5	882	85.6	882	4	US-10-072-012-622	Sequence 622, Appl
6	882	85.6	882	4	US-10-415-122-6	Sequence 6, Appli
7	882	85.6	882	4	US-10-825-633-1	Sequence 1, Appli
8	882	85.6	882	5	US-10-982-512-1	Sequence 1, Appli
9	655	63.6	658	3	US-09-976-674-19	Sequence 19, Appl
10	655	63.6	658	5	US-10-982-512-19	Sequence 19, Appl
11	655	63.6	661	3	US-09-976-674-11	Sequence 11, Appl

ALIGNMENTS

```

RECORD 1
US-09-976-674-1
; Sequence 1, Application US/09976674
; Patent No. US20020115843A1
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 882
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-1

```

Db 1 MetAlaAlaMetGluThrGluGluLeuGlyValGluIlePheGluThrAlaAspCys 20
Qy 274 GAGGAGATATTGAATCACAGATCGGCCTTAATTTGGAGCCCTTTATTTGTTGAGCGGTAT 333
Db 21 GluGluAenIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
Qy 334 TCTGGAGTCAGCTTAAAGCTTGGCCGATACAGAAAATATATCATGGCTACATGATG 393
Db 41 SerTrpSerGlnLeuLysLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
Qy 394 GCTAAGGCCACCATGATTTTCATGTTTGTGAAGAGAAATGATCCAGATGGACCTCATCA 453
Db 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
Qy 454 GACAGAATCTATTACCTTGGCATGTCTGTGTGAGAACAGAGAAAATACACTGTTTATTCT 513
Db 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
Qy 514 GAAATTTCCAAAACATATCAATAGACGAGCAGTCTTAATGCTCTCTTGGAAAGCCTTTTG 573
Db 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120
Qy 574 GATCTTTTTCAGCAACACACTGGACTATGGAATGTATTCTCGAGAGAAAGAACTATTAGA 633
Db 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg 140
Qy 634 GAAAGAAACCATTTGAAACAGTCGGAATTTGCTTACAGATTATCCACAGGAGAGTGA 693
Db 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
Qy 694 ACNTTTCNTGTTCAAGCCGCTAGTGGAAATTTATCAGTAAAGATGGAGGCCACAGGA 753
Db 161 ThrPheLeuPheGlnAlaGlySerGlyTyrHisValLysAspGlyGlyProGlnGly 180
Qy 754 TTTACGCCAACCTTTAAGGCCCAATCTAGTGGAAACTAGTTGTGCCCAACATACGGATG 813
Db 181 PheThrGlnGlnProLeuLysProAsnLeuValGluThrSerCysProAsnIleArgMet 200
Qy 814 GATCCAAATTTATCCCGCTGATCCAGACTGGATGCTTTTATATACATAGCAACGATATT 873
Db 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
Qy 874 TGGATATCTAAACATCGTAACACAGAGAAAGAGAGACTCATTTATGTCCACAAATGAGCTA 933
Db 221 TrpIleSerAsnIleValThrArgGluArgLysProLysAlaGluThrThrProSerGlyGly 240
Qy 934 GCCAACATGGGAAGATGCCAGATCAGCTGGAGTCGCTTACCTTTTGTCTCCAAAGAAAGAA 993
Db 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGluGluGlu 260
Qy 994 TTTGATAGATATTCTGGCTATTGTTGTTGTCAAAAGCTGAAACAACTCCCGAGTGGTGT 1053
Db 261 PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly 280
Qy 1054 AAAATTTCTAGAATTTCTATATGAAGAAATGATGATCTGAGTGGAAATTTATTCATGTT 1113
Db 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal 300
Qy 1114 ACATCCCTATGTTGGAACACAGAGCGCAGATTCAATTCCTGTTATCTTAAACAGGTACA 1173
Db 301 ThrSerPrometLeuGluThrArgAlaAspSerPheArgTyrProLysTrpGlyThr 320
Qy 1174 GCAATCTTAAGTCATTTTAAGATGTCAGAAATATGATTTGATGCTGAAGGAGGATC 1233
Db 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
Qy 1234 ATAGATGTCATAGTAAGAACCTAATTCACCTTTTGTAGATCTTATTTGAAGGAGTTGAA 1293
Db 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
Qy 1294 TATATTGCCAGAGCTGGAATCGACTCTCGAGGAGAAATATGCTTGTGTCATCTACTAGAT 1353
Db 361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380

Qy 1354 CGCTCCAGACTCGCCTACAGATAGTGTGTTGATCTCACTGAATATTATTTATCCAGTAGAA 1413
Db 381 ArgSerGlnThrArgLeuGluIleValLeuIleSerProGluLeuPheIleProValGlu 400
Qy 1414 GATGATGTTATGGAAAGGCAGAGACTCATGTAGTCCAGTCCCTGATCTTGTGACGCCACTA 1473
Db 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
Qy 1474 ATTATCTATGAAGAAACACACAGACATCTCGATAAATATCCATGACATCTTTCATGTTTT 1533
Db 421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
Qy 1534 CCCCAAGATCACGAAGAGGAAATTTGATTTTATTTTTCCTCTGAATTCGAAACAGGTTTC 1593
Db 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
Qy 1594 CGTCATTTATACAAAATTTACATCTATTTTAAAGGAAAGCAAAATATAAACGATCCAGTGGT 1653
Db 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
Qy 1654 GGGCTGCTGCTCCCAAGTGTCTCAAGTGTCTCTATCAAAGAGAGAGATAGCAATTTACCAGT 1713
Db 481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluIleAlaIleThrSer 500
Qy 1714 GGTCAATGGGAAGTCTTGGCCGCATGGATCTAATATCCAGTTTGTAGTGAAGTCAGAGG 1773
Db 501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520
Qy 1774 CTGATATATTTGAAGCCACCAAGACTCCCTTTTAGACATCACTCTAGTGTAGT 1833
Db 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSer 540
Qy 1834 TACGTAAATCTCGAGAGGTGACAGCTGACTGACCGTGGCTACTCATCTTCTGCTGC 1893
Db 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
Qy 1894 ATCAGTCAGACTGTGACTCTTTTATAGTAAGTATAGTAAACAGAGAAATCCACACTGT 1953
Db 561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
Qy 1954 GTGTCCCTTTTCAAGCTATCAAGTCTCTGAAGATGACCCCACTTGCAAAACAAAGAAATTT 2013
Db 581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
Qy 2014 TGGCCCACTTTTGGATTCAGCAGGTCCTCTTCTCGACTATCTCTCCAGAAATTTTC 2073
Db 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe 620
Qy 2074 TCTTTTCAAAGTACTACTGGATTTACATTTGATGGATGCTCTCAAGCCCTCATGATCTA 2133
Db 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
Qy 2134 CAGCTCGAAAGAAATATCTACTGTGCTGTTTCATATATGTTGGTGGTCTCAGGTGAGTTG 2193
Db 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu 660
Qy 2194 GTGAATTAATCGGTTTAAAGGAGTCAGTATTTCCCTTGAATACCTAGCTCTCTAGGT 2253
Db 661 ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly 680
Qy 2254 TATGTGTTGTAGTATAGACAACAGGCGATCTCTCCACCGGGCTTAAATTTGAAGGC 2313
Db 681 TyrValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly 700
Qy 2314 GCCTTTAAATATATAAATGGGTCAATAGAAATTTGACATCAGTCCAGTGGAGGACTCAATAT 2373
Db 701 AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr 720
Qy 2374 CTAGCTTCTCGATATGATTTTCATTTGATCTAGTTCGTGTGGGCATCCAGGCTGCTCTAT 2433
Db 721 LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyr 740

QY 2434 GAGGATACCTCCCTGATGCATTAAATGACAGAGGTTCAGATATCTTCAGGTTGCTATT 2493
Db |||||
QY 741 GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760
Db |||||
QY 2494 GCTGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACAGGAACTTATATG 2553
Db |||||
QY 761 AlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluA-GTyrMet 780
Db |||||
QY 2554 GGTCAACCTGACAGAAATGAACAGGGCTATTACTTATGATCGGATCGGCAAGCAGAA 2613
Db |||||
QY 781 GlyHisProAspGlnAsnGlnGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu 800
Db |||||
QY 2614 AGTTCCCTCTGAACCAATCGTTTACTGCTCTTACATGCTTTCCTGGATGATGTC 2673
Db |||||
QY 801 LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820
Db |||||
QY 2674 CATTTTGCACATACCAAGTATATCTAGTGTCTTGTAGTGAGGCTGGAAAGCCATATGAT 2733
Db |||||
QY 821 HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840
Db |||||
QY 2734 TTACAGATCTATCTCAGGAGACACAGACATAAGAGTTCTCTGAATCGGAGAACATAT 2793
Db |||||
QY 841 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860
Db |||||
QY 2794 GAACCTGATCTTTTGCACACTACCTTCAAGAAACCTTTGGATCACGTATTGTGCTCTTAAA 2853
Db |||||
QY 861 GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeuLys 880
Db |||||
QY 2854 GTGATA 2859
Db |||||
QY 881 ValIle 882

RESULT 2
US-10-054-776-2
; Sequence 2, Application US/10054776
; Publication No. US20030165818A1
; GENERAL INFORMATION:
; APPLICANT: Mark Robert Edbrooke
; APPLICANT: Alan Peter Lewis
; TITLE OF INVENTION: NOVEL PROTEIN
; FILE REFERENCE: QG1042US
; CURRENT APPLICATION NUMBER: US/10/054,776
; CURRENT FILING DATE: 2002-01-23
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 882
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-054-776-2

Alignment Scores:
Pred. No.: 0 Length: 882
Score: 882.00 Matches: 882
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 85.6% Indels: 0
DB: 4 Gaps: 0

US-10-825-632-2 (1-3120) x US-10-054-776-2 (1-882)

QY 214 ATGGCAGCAGCAATGGAACAGAACAGAGCTGGGTGTGAGATATTTGAAACTGCGAGTGT 273
Db |||||
QY 1 MetAlaAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
Db |||||
QY 274 GAGGAGAAATATGAAATCACAGGATCGGCCTTAAATTTGGAGCCCTTTTATGTTGAGCGGTAT 333
Db |||||
QY 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
Db |||||
QY 334 TCCTGGATGAGCTTAAAGCTGCTTGGCGATACAGAAATATCATGCTACATGATG 393
Db |||||
QY 41 SerTrpSerGlnLeuLysLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
Db |||||

QY 394 GCTAAGGCCACCATCATTTTCATGTTTCTGAAGAGGAATGATCCAGATGGACCTCATTTCA 453
Db |||||
QY 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
Db |||||
QY 454 GACAGAATCTATTACCTTCCCATGCTCGTGTGAGAACAGAGAAATACACTGTTTATTCT 513
Db |||||
QY 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
Db |||||
QY 514 GAAATTTCCCAAACTATCAATAGACGACGAGCTTAATGCTCTCTCGAAGCCCTCTTTTG 573
Db |||||
QY 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120
Db |||||
QY 574 GATCTTTTTCAGGCAACACTGACTATGGAATGTATTCTCGAGAGAAAGAACTATTAAAGA 633
Db |||||
QY 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg 140
Db |||||
QY 634 GAAAGAAAACGCAATTTGGAACAGTCGGAATGCTTCTTACGATTATTCACAAAGGAAGTGA 693
Db |||||
QY 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
Db |||||
QY 694 ACATTTCTGTTTCAAGCCGCTAGTGGATTTTATTCAGTAAAAAGATGGAGGCCCAAGGA 753
Db |||||
QY 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
Db |||||
QY 754 TTTACGCAACAACTTTAAGGCCCACTCTAGTGGAACTAGTTGTCCCAACATACGATG 813
Db |||||
QY 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
Db |||||
QY 814 GATCCAAAATTATGCCCCCTGATCCAGACTCGGATGCTTTTATACATAGCAACGATATT 873
Db |||||
QY 201 AspProLysLeuCysProAlaAspProAspThrIleAlaPheIleHisSerAsnAspIle 220
Db |||||
QY 874 TGGATATCTAACATCGTAACACAGAGAGAAAGAGAGACTCCTTATGTGCACAATGAGCTA 933
Db |||||
QY 221 TrpIleSerAsnIleValThrArgGluArgGluLeuThrTyrValHisAsnGluLeu 240
Db |||||
QY 934 GCCAACATGGAAGAAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAGAGAA 993
Db |||||
QY 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu 260
Db |||||
QY 994 TTTGATAGATATTCTGGCTATTGTTGGTGTCCAAAAGCTGAAACAACTCCCAAGTGGTGT 1053
Db |||||
QY 261 PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly 280
Db |||||
QY 1054 AAAATTCTTGAATTTCTATATGAAGAAAATGATGAATCTGAGGTGGAATTTATTCATGTT 1113
Db |||||
QY 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal 300
Db |||||
QY 1114 ACATCCCTCTATGTTGGAACACAGAGGCGAGATTCAATCCGTTATCTCTAAACAGGTACA 1173
Db |||||
QY 301 ThrSerProMetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThr 320
Db |||||
QY 1174 GCAAACTCTAAAGTCACTTTTAAAGATGTGAGAAATAATGATGTGCTCCATCTACTAGAT 1233
Db |||||
QY 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
Db |||||
QY 1234 ATAGATGTCATAGTAAGGAACCTAAATCAACCTTTTGAGATTCTATTGAGGAGGTCAA 1293
Db |||||
QY 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
Db |||||
QY 1294 TATATTTGCCAGAGCTGATCGACTCTCTGAGGGAATAATATGCTTGGTCCATCTACTAGAT 1353
Db |||||
QY 361 TyrIleAlaArgAlaGlyTyrThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
Db |||||
QY 1354 CGCTCCACAGACTCCGCTTACAGATAGTGTGATCTCACTCGAATATTTATCCAGTAGAA 1413
Db |||||
QY 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
Db |||||
QY 1414 GATGATGTTATGGAAGGCGAGAGACTCATGTAGTCGCTGATTCGTGAGCGCCTA 1473
Db |||||
QY 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
Db |||||
QY 1474 ATTATCTATGAAGAAACACAGACATCTGATAAATATCCATGACATCTTTTCATGTTTT 1533
Db |||||

Db 421 IleIleTyrGluGluThrThrAspIleThrPheAenIleHisAspIlePheHisValPhe 440
QY 1534 CCCAAAGTCACGAAGAGGAAATGAGTTATTTTTCCTCTGAATGCAAAACAGGTTTC 1593
Db 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
QY 1594 GGTCAATTTATACAAATTTACATCTATTTTAAAGGAAGCAATATATAAGAGTCCAGTGGT 1653
Db 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysIleThrSerGly 480
QY 1654 GGGCTGCTGCTCCCAAGTGATTTCAAGTGTCTCTATCAAGAGGAGATAGCAATACCAGT 1713
Db 481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer 500
QY 1714 GGTGAATGGGAAGTCTTGGCCGGCATGGATCTATATCAAGTTGTATGAAGTCAGAAGG 1773
Db 501 GlyGluThrGluValLeuGlyArgHisGlySerAenIleGlnValAspGluValArgArg 520
QY 1774 CTGGTATATTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACCTGTACGTAGTCAGT 1833
Db 521 LeuValTyrPheGluGluThrLysAspSerProLeuGluHisLeuTyrValValSer 540
QY 1834 TACGTAAATCTGGAGAGGTGACAAAGGCTGACTGACCGTGGCTACTCACATTTCTTGCTGC 1893
Db 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
QY 1894 ATCAGTCAGCACTGTGACTTCTTTATAAGTAAGTATAGTAACCAAGAAATCCACACTGT 1953
Db 561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
QY 1954 GTGTCCCTTTACAAAGCTATCAGTCTCTGAAGATGACCAACTTCGCAAAACAAAGGAATTT 2013
Db 581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
QY 2014 TGGGCCACCAATTTGATTGATTCAGCAGGTCTCTCTCTGACTATCTCTCCAGAAATTTTC 2073
Db 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIlePhe 620
QY 2074 TCTTTTGAAGTACTACTCGATTTACATGTATGGATGGTCTCAAGCTCATGATCTA 2133
Db 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
QY 2134 CAGCTGGGAAGAAATCTCTACTGCTGCTTCATATATGGTGTCTCAGGTGCAAGTTG 2193
Db 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu 660
QY 2194 GTCAATAACGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCCTAGCCTCTCTAGT 2253
Db 661 ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly 680
QY 2254 TATGTGGTTGTGTAGATAGACAAACAGGGGATCCTGTCAACGAGGGCTTAAATTTGAAGGC 2313
Db 681 TyrValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly 700
QY 2314 GCCTTTAAATATAAAATGGGTCAAAATAGAAATTCACGATCAGGTGGAAGGACTCCCAATAT 2373
Db 701 AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr 720
QY 2374 CTAGCTTCTCGATATGATTTCAATTGACTTAGATCGTGGGCATCCACGGCTGGCTCAT 2433
Db 721 LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyr 740
QY 2434 GGAGGATACCTCTCCCTGATGTCATTATGAGAGGTACAGATATCTTCAGGGTTGCTATT 2493
Db 741 GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760
QY 2494 GTGGGGCCCCAGTCACCTGTGGATCTTCTATGATACAGGATACACGGAAGGTTATATG 2553
Db 761 AlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMet 780
QY 2554 GTTCACCTCGACCAAGTGAACAGGGCTATTACTAGGATCTGTGGCCATGCAACAGAA 2613

Db 781 GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetClnAlaGlu 800
QY 2614 AAGTTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGGTTTCTCGATGAGAATGTC 2673
Db 801 LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820
QY 2674 CATTTTGCACATACCACTAGTATATTTACTGAGTCTTTTGTAGTGAGGCTCGAAAGCCATATGAT 2733
Db 821 HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840
QY 2734 TTACAGATCTATCTCTCAGGAGAGACACAGATTAAGAGTTTCTCTGAATCGGAGAAACATTAT 2793
Db 841 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860
QY 2794 GAACTGCATCTTTTGACACTACCTTCAAGAAACCTTGGATCACGTATTGCTGCTCTAAA 2853
Db 861 GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaLeuLys 880
QY 2854 GTGATA 2859
Db 881 ValIle 882
RESULT 3
US-10-170-789-38
; Sequence 38 Application US/10170789
; Publication No. US20030180930A1
; GENERAL INFORMATION:
; APPLICANT: Rachel E. Meyers
; APPLICANT: Olandt, Peter J.
; APPLICANT: Kapeller-Ribermann, Rosana
; APPLICANT: Cuthrie, Rory A. J.
; APPLICANT: Williamson, Mark
; APPLICANT: Meich, Nadine
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASE, PHOSPHATASE,
; TITLE OF INVENTION: AND PROTEASE FAMILY MEMBERS AND USES THEREOF
; FILE REFERENCE: 10448-191001
; CURRENT APPLICATION NUMBER: US/10/170,789
; CURRENT FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: US 09/797,039
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: PCT/US01/06525
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/186,061
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 09/882,166
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: PCT/US01/19269
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/212,078
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 09/934,406
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: PCT/US01/26052
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/226,740
; PRIOR FILING DATE: 2000-08-21
; PRIOR APPLICATION NUMBER: US 09/861,801
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: PCT/US01/16549
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: US 60/205,508
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: US 09/801,267
; PRIOR FILING DATE: 2001-03-06
; PRIOR APPLICATION NUMBER: PCT/US01/07138
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: US 60/187,454
; PRIOR FILING DATE: 2000-03-07
; PRIOR APPLICATION NUMBER: US 09/829,671
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: PCT/US01/40483
; PRIOR FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: US 60/197,508

```
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: US 09/961,721
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: PCT/US01/29904
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: US 60/235,023
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US 10/045,367
; PRIOR FILING DATE: 2001-11-07
; PRIOR APPLICATION NUMBER: US 60/246,561
; PRIOR FILING DATE: 2000-11-07
; PRIOR APPLICATION NUMBER: US 09/801,275
; PRIOR FILING DATE: 2001-03-06
; PRIOR APPLICATION NUMBER: PCT/US01/07074
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: US 60/187,420
; PRIOR FILING DATE: 2000-03-07
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 38
; LENGTH: 882
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-170-789-38

Alignment Scores:
Pred. No.: 0 Length: 882
Score: 882.00 Matches: 882
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 85.6% Indels: 0
DB: 4 Gaps: 0

US-10-825-632-2 (1-3120) x US-10-170-789-38 (1-882)

QY 214 ATGCGACGACAAATGAAACAGACAGCTGGTGTGTGAGATATTGAAACATGGCGACTGT 273
Db 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
QY 274 GAGGAGAATATTGAATCAGAGATCGGCTAAATTTGGAGCCCTTTTATTTGTTGACCGTAT 333
Db 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
QY 334 TCCTGGAGTCAGCTTAAAGCTGCTGGCGATACAGAAATATCATGGCTCATCATG 393
Db 41 SerTrpSerGlnLeuLysLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
QY 394 GCTAAGCACACATGATTTTCATGTTGTGAAGAGAAATGATCCAGATGGACCTCATTTCA 453
Db 61 AlalysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
QY 454 GACAGAATCTATTACCTTGCATGCTCGGTGAGAACAGACAGAGAAATACACTGTTTATTTCT 513
Db 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
QY 514 GAAATTTCCCAAACTATCAATAGACGACGAGCTTTAATGCTCTCTTGGAAACCTCTTTTG 573
Db 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120
QY 574 CATCTTTTTCAGGCAACACTGGACTATGGAATGTATTCTCGAGAGAGAGAACTATTAGA 633
Db 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg 140
QY 634 GAAAGAAACCATTTGAAACAGTCGGAATTTGCTTCTTACGATTATCCCAAGGAGGTGA 693
Db 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
QY 694 ACATTTCTGTTTCAAGCCGCTAGTGGAAATTTATTCAGTAAAGATGGAGGGCCCAAGGA 753
Db 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
QY 754 TTTACGCAACACCTTTTAAGGCCCAATCTAGTGGAACTAGTTGTCCTCCCAACATACGGATG 813
Db 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
QY 814 GATCCAAAATTTATGCCCGCTGATCCAGACTGGGATTGCTTTTATACATAGCAACGATATT 873
Db 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
QY 874 TGGATATCTAACATCGTTAAACAGAGAGAAAGGAGACTCACTTATGTGCAATAGACTA 933
Db 221 TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu 240
QY 934 GCCAACATGGAAGAGATGCCAGATCAGCTGAGTGCCTACCTTTGTTCTCCCAAGAGAA 993
Db 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu 260
QY 994 TTTGATAGATATTCTGGCTATTGGTGTCTCCAAAGCTGAAACAACTCCACAGTGGTGGT 1053
Db 261 PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly 280
QY 1054 AAAATTCTTAGAATCTATATGAAGAAATGATGAATCTGAGGTGGAAATTTATTCATGTT 1113
Db 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleHisVal 300
QY 1114 ACATCCCCTATGTTGAAACAAAGGAGGCGAGATTCATTCCTGTTATCTCTAAACAGGTACA 1173
Db 301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr 320
QY 1174 GCAATCTCTAAAGTCACTTTTAAAGATGTCAGAAATATGATGATGATGTCGTAAGGAGGATC 1233
Db 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
QY 1234 ATAGATGTCATAGATAAGAACTAATTAACCTTTTGGAGATCTTATTTGAGGAGGTGAA 1293
Db 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
QY 1294 TATATTGCCAGACTCGATGGACTCTCGAGGGAATATGCTTGCTCCTACTACTAGAT 1353
Db 361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
QY 1354 CGCTCCCGACTCGCTACAGATAGTGTGATCTCACCTGAAATTTATTTATCCAGTAGAA 1413
Db 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
QY 1414 GATGATGTTATGAAAGGAGAGACTCAITGAGTCAGTGCCTGATCTGTGACGCCACTA 1473
Db 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
QY 1474 ATTATCTATGAAGAAACACAGACATCTGGATAAATATCCATGACATCTTTCATGTTTTT 1533
Db 421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
QY 1534 CCCCAGAGTCACGAGGAGAAATGAGTTATTTTGGCTCTGATGATGCAAAACAGGTTTC 1593
Db 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
QY 1594 CGTCATTTATACAAAATTTACATCTATTTTAAAGGAAAGCAATATATAACGATCCAGTGGT 1653
Db 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
QY 1654 GGGCTGCTGCTCCAAAGTGAATTTCAAGTGTCTCTATCAAGAGGAGATAGCAATTTACCAGT 1713
Db 481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer 500
QY 1714 GGTGAATGGGAAGTCTTTGGCCCGCATGATCTAATATCCAAAGTTGATGAAGTCAGAGG 1773
Db 501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520
QY 1774 CTGGTATATTTTGAAGGCCAACAGACTCCCTTTTAGAGCATCACCTGTACGTACTGTCAGT 1833
Db 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSer 540
QY 1834 TACCTTAATCTCGAGAGGTCACAGCGTACCTGACCGTGGCTACTCTACATTTCTTGCTGC 1893
Db 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
```

```
QY 1894 ATCAGTCAGCACTGTGACTCTCTTTATTAAGTAAGTATAGTAACCAAGAAATCCACACTGT 1953
Db 561 lIleSerGlnHisCysAaspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
QY 1954 GTGTCCCTTTTACAAAGCTATCAAGTCCTGAAGATGACCCAACTTGCABAACAAAGAAATTT 2013
Db 581 ValSerLeuTyrLysLeuSerSerProGluAaspAapProThrCysGlyThrLysGluPhe 600
QY 2014 TGGGCCACCAATTTTGGATTCAGCAGGTCTCTCTCTGACTATCTCTCCAGAAATTTTC 2073
Db 601 TrpAlaThrIleLeuAaspSerAlaGlyProLeuProAaspTyrThrProProGluIlePhe 620
QY 2074 TCTTTTGAAGTACTACTGATTTACATTGTATGGGATGCTCTCAAGCCCTCATGATCTA 2133
Db 621 SerPheGluSerThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAaspLeu 640
QY 2134 CAGCCTCGAAGAAATATCTACTGTGCTGTTCAATATATGGGFCCTCAGGTGCAGTTG 2193
Db 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu 660
QY 2194 GTGAATATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCCCTAGCCTCTCTAGGT 2253
Db 661 ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAenThrLeuAlaSerLeuGly 680
QY 2254 TATGTGTTGTAGTATAGAACACAGGGGATCTCTGTCCAGCGGCTTAAATTTGAAGGC 2313
Db 681 TyrValValValLysMetGlyGlnIleGluIleAaspAaspGlnValGluGlyLeuGly 700
QY 2314 GCCTTTAAATATAAAATGGGTCAAATAGAAATTGACGATCAGGTGGAAGGACTCCAATAT 2373
Db 701 AlaPheLysTyrLysMetGlyGlnIleGluIleAaspAaspGlnValGluGlyLeuGlnTyr 720
QY 2374 CTAGCTTCTCGATATGATTTCAATTCATCTAGATCTGTGGGCATCCAGCTGGTCTAT 2433
Db 721 LeuAlaSerArgTyrAspPheIleAaspLeuAaspValGlyIleHisGlyTrpSerTyr 740
QY 2434 GGAGGATACCTCTCCCTGATGCATTAATGACAGGTCAGATATCTTCAGGTTGCTATT 2493
Db 741 GlyGlyTyrLeuSerLeuMetAlaMetGlnArgSerAaspIlePheArgValAlaIle 760
QY 2494 GCTGGGCCCGAGTCACTCTGGGATCTCTTCTATGATACAGATACAGGAACGTTATATG 2553
Db 761 AlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMet 780
QY 2554 GGTACCGCTGACCAGAAATGAAACAGGCTATTACTTAGATCTGTGGCCATGCAAGCAGAA 2613
Db 781 GlyHisProAaspGlnAsnGlnGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu 800
QY 2614 AAGTCCCTCTGAAACCAAAATCGTTTACTGCTCTTACATGCTTTCTGGATGAGAATGTC 2673
Db 801 LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAaspGluAenVal 820
QY 2674 CATTTTGCATACACAGTATATTACTGAGTGTCTTTTAGTGAGGCTGGAAAGCCATATGAT 2733
Db 821 HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAap 840
QY 2734 TTACAGATCTATCCTCAGGAGACACAGCATAAGAGTCTCTGAATCGGAGAACATTAAT 2793
Db 841 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860
QY 2794 GAACTCATCTTTTGCACCTACCTTCAAGAAACCTTCGATCAGTATCTCTCTCTCAAA 2853
Db 861 GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaLeuLys 880
QY 2854 GTGATA 2859
Db 881 ValIle 882
```

RESULT 4
US-10-311-035-9
; Sequence 9, Application US/10311035
; Publication No. US20040023243A1

```
GENERAL INFORMATION:  
; APPLICANT: INCYTE GENOMICS, INC.  
; APPLICANT: YUE, Henry  
; APPLICANT: ELLIOTT, Vicki  
; APPLICANT: GANDHI, Ameena R.  
; APPLICANT: LAL, Preeti  
; APPLICANT: AU-YOUNG, Janice  
; APPLICANT: TRIBOULEY, Catherine M.  
; APPLICANT: DELEGEANE, Angelo M.  
; APPLICANT: BAUGHN, Mariah R.  
; APPLICANT: NGUYEN, Dannel B.  
; APPLICANT: LEE, Ernestine A.  
; APPLICANT: HAFALIA, April  
; APPLICANT: KHAN, Farrah A.  
; APPLICANT: CHAWLA, Narinder K.  
; APPLICANT: YAO, Monique G.  
; APPLICANT: LU, Dyung Aina M.  
; APPLICANT: ARVIZU, Chandra S.  
; APPLICANT: TANG, Y. Tom  
; APPLICANT: WALSH, Roderick T.  
; APPLICANT: AZIMZAI, Valda  
; APPLICANT: LU, Yan  
; APPLICANT: RAMKUMAR, Jayalaximi  
; APPLICANT: XU, Yuming  
; APPLICANT: REDDY, Roopa  
; APPLICANT: DAS, Depopriya  
; APPLICANT: KERANEY, Piam  
; APPLICANT: KALLICK, Deborah A.  
; TITLE OF INVENTION: Processes  
; FILE REFERENCE: FI-0123 PCT  
; CURRENT APPLICATION NUMBER: US/10/311,035  
; PRIOR FILING DATE: 2002-12-10  
; PRIOR APPLICATION NUMBER: 60/212,336; 60/213,955; 60/215,396; 60/216,821; 60/218,946  
; PRIOR FILING DATE: 2000-06-16; 2000-06-22; 2000-06-29; 2000-07-07; 2000-07-14  
; NUMBER OF SEQ ID NOS: 42  
; SOFTWARE: PERL Program  
; SEQ ID NO 9  
; LENGTH: 882  
; TYPE: PR1  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc_feature  
; OTHER INFORMATION: Incyte ID No. US20040023243A1 7160544CD1  
US-10-311-035-9  
  
Alignment Scores:  
Pred. No.: 0 Length: 882  
Score: 882.00 Matches: 882  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 85.6% Indels: 0  
DB: Gaps: 0  
  
US-10-825-632-2 (1-3120) x US-10-311-035-9 (1-882)  
QY 214 ATGGCAGCAGCAATCGAAACAGACAGCTGGCTTCAGATATTTGAACTCGGACTGT 273  
Db 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAaspCys 20  
QY 274 GAGGAGATATTGATCATCAGATCGGCCTTAAATTGAGGCTTTTATGTTGAGCGGTAT 333  
Db 21 GluGluAenIleGluSerGlnAaspProLysLeuGluProPheTyrValGluArgTyr 40  
QY 334 TCTGTGAGTCAGCTTAAAAAGCTGCTTCGCGATACCAAGAAATATCATGCTACATGATG 393  
Db 41 SerTrpSerGlnLeuLysLeuLeuAlaAaspThrArgLysTyrHisGlyTyrMetMet 60  
QY 394 GCTAAGGCACCAACATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGGACCTCATTC 453  
Db 61 AlaLysAlaProHisAaspPheMetPheValLysArgAsnAaspProAaspGlyProHisSer 80  
QY 454 GACAGAATCTATTACCTTCCCATCTCTGGTGAGAACAGAGAAATACACTGTTTATTCT 513
```

Db 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
QY 514 GAAATCCCAAACTATCAATAGACAGCAGTCTTAATGCTCTCTGGAAGCCTCTTTG 573
Db 101 GluIleProLysThrIleAsnArgAlaAlaValIleMetLeuSerTrpLysProLeuLeu 120
QY 574 GATCTTTTTCAGGCAACACATCGACTATGGAAATGATTTCTCGAGAAGAACTATTAGA 633
Db 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg 140
QY 634 GAAAGAAACCATTTGGAACAGTCGGAATTCCTTACGATTATCAACCAAGGAAGTGA 693
Db 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
QY 694 ACATTTCTGTTTCAAGCGGTAGTGGAAATTTATCAGTAAAGATGGAGGGCCACAAGA 753
Db 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
QY 754 TTTAGCCCAACACCTTTAAAGCCCAATCTAGTGGAACTAGTTGCTGCCAACAATAGGATG 813
Db 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
QY 814 GATCCAAATTAATGCCCCGTGATCCAGACTGGATTGCTTTTATATACATAGCAACGATATT 873
Db 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
QY 874 TGGATATCTAAATCGTAACACAGAGAAGAGAGACTCACTATGTGCAACATGAGCTA 933
Db 221 TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu 240
QY 934 GCCAACATGGAAGATGCCAGATCAGCTGAGTCGCTACCTTTGTTCTCCAGAAGAA 993
Db 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu 260
QY 994 TTTGATAGATATTCGGCTATTTGGTGTGCCAAAAGCTGAAACAACTCCCAAGTGGTGGT 1053
Db 261 PheAspArgTyrSerGlyTyrTrpTyrCysProLysAlaGluThrThrProSerGlyGly 280
QY 1054 AAAATCTTAGAATCTATATGAGAANAATGATGATCTGAGGTGGGAATTTATCATGTT 1113
Db 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleHisVal 300
QY 1114 ACATCCCTATGTTGGAACAAGAGGGCAGATTCATTCGGTTATCCCTAAACAGGTACA 1173
Db 301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr 320
QY 1174 GGAATCTCTAAAGTCACCTTTTAAGATGTCAGAAATAATGATGATGCTGGAAGAGGATC 1233
Db 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
QY 1234 ATAGATGTCATAGATAAGGAACATAATCAACCTTTTGGAGATCTATTGGAAGGATTGAA 1293
Db 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
QY 1294 TATATGTCAGAGCTGGATGATCCTCTGAGGGAATAATGCTTGGTGCATCTCTAGAT 1353
Db 361 TyrIleAlaArgAlaGlyTyrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
QY 1354 CGCTCCAGACTCCGCTACAGATAGTGTGATCTACCTCGAATTTATCCAGTAGAA 1413
Db 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
QY 1414 GATGATGTTATGGAAGGCAGACTCATTTGACTGAGTCGCTGATTCGTGACGCCACTA 1473
Db 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
QY 1474 ATTATCTATGAAGAAACAACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTTT 1533
Db 421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
QY 1534 CCCCAGTCAAGAGGAATAGTTTTATTTTTTTCCTCTGAAATGCAAAACAGGTTTC 1593
Db 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460

QY 1594 CGTCATTTTATACAAAATTACATCTTATTTTAAAGAAAGCAAATATATAACGATCCAGTGGT 1653
Db 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
QY 1654 GGGCTGCTGCTCCCAAGTGATTTTCAAGTGCTCTCAAAAGAGAGAGATAGCAATTAACAGT 1713
Db 481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer 500
QY 1714 GGTGAATGGGAAGTCTTTGCGCGCATGGAATCTAATATCAAGTTGATGAAGTCAAGAG 1773
Db 501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArg 520
QY 1774 CTGGTATATTTTGAAGCACCAAGACTCCCTTTAGACATCACCCTGTACCTAGTCACT 1833
Db 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisIleLeuTyrValValSer 540
QY 1834 TACGTAAATCTCGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACATTTCTTGCTGC 1893
Db 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
QY 1894 ATCAGTCAGCACTGTGACTTCTTTTATAAGTAAGTATAGTAACCAAGAAATCCACACTGT 1953
Db 561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
QY 1954 GTGTCCTTTTCAAGCTATCAAGTCTCTGAAGATGACCCCACTTGCAAAAACAAAGAAATTT 2013
Db 581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
QY 2014 TGGGCCACCATTTTGGATTCAGCAGGTCTCTTCTCTGACTATCTCTCCAGAAATTTTC 2073
Db 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIlePhe 620
QY 2074 TCTTTTGAAGTACTACTGGATTTACATATGATGGGATGCTCTACAGCCTCATCATCTA 2133
Db 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
QY 2134 CAGCTCGAAAAGAAATCTCTACTGTCTTTCATATATGTTGGTCTCAGGTGCAGTTG 2193
Db 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyProGlnValGlnLeu 660
QY 2194 GTGAATATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTAGCCTCTCTAGT 2253
Db 661 ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly 680
QY 2254 TATGTGGTGTGATGATAGACAACAGGGATCTCTGTCCACGGGGCTTAAATTTGAAGGC 2313
Db 681 TyrValValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly 700
QY 2314 GCCTTTAAATATAAATGGGTCAAAATAGAAATTTGACGATCAGGTGGAAGGACTCCAATAT 2373
Db 701 AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr 720
QY 2374 CTAGCTTCTCATATGATTTTTCATTTGACTTAGATCGTGGGCATCCACGGCTGCTCAT 2433
Db 721 LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyr 740
QY 2434 GGAGGATACCTCTCCCTGATGGCATTAATGACAGGTACAGATATCTTCCAGGTTGCTATT 2493
Db 741 GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760
QY 2494 GCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAACGTTATATG 2553
Db 761 AlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMet 780
QY 2554 GGTCAACCTGACAGAAATGAACAGGGCTATTACTAGGATCTCTGGCCATGCAACAGAA 2613
Db 781 GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu 800
QY 2614 AAGTTCCCTCTGAACCAATCGTTTACTGCTCTTACATGTTTCTTGATGAGATGTC 2673
Db 801 LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820


```
QY 2674 CATTTCGACATACAGTATATTAAGTCTTTTGTAGTGGCTGGAAAGCCATATCAT 2733
Db 821 HISPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840

QY 2734 TTACAGATCTATCTCTAGGAGACACAGCATTAAGATTCCTGAATCGGAGAACATAT 2793
Db 841 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860

QY 2794 GAATCGCATCTTTTGCACCTACCTTCAAGAAACCTTTGGATCACGATTGCTGCTCAAAA 2853
Db 861 GluLeuHisLeuLeuHisTyrLeuGlnGluAenLeuGlySerArgIleAlaAlaLeuLys 880

QY 2854 GTGATA 2859
Db 881 ValIle 882

RESULT 5
US-10-072-622
; Sequence 622, Application US/10072012
; Publication No. US20040033493A1
; GENERAL INFORMATION:
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spyrek, Kimberly
; APPLICANT: Zerkusen, Bryan
; APPLICANT: Patturajan, Meera
; APPLICANT: Shinkets, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangolli, Esha
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Taupier Jr, Raymond J.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Colman, Steven D.
; APPLICANT: Wolenc, Adam R.
; APPLICANT: Pena, Carol E. A
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Grosse, William M.
; APPLICANT: Alsbrook II, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-258
; CURRENT APPLICATION NUMBER: US/10/072,012
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/265,102
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265,514
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/266,406
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/266,767
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/267,057
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/266,975
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/267,459
; PRIOR FILING DATE: 2001-02-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1391
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 622
; LENGTH: 882
```

```
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-072-012-622

Alignment Scores:
Pred. No.: 0 Length: 882
Score: 882.00 Matches: 882
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 85.6% Indels: 0
DB: 4 Gaps: 0

US-10-825-632-2 (1-3120) x US-10-072-012-622 (1-882)

QY 214 ATGGCAGCAGCAATGGAACACAGAACAGCTGGGTGTGTGAGATATTTGAAACTGCGACTGT 273
Db 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20

QY 274 GAGGAGAATATTGAATCACAGGATCGCCTAAATTTGGAGCCTTTTATTTATTTGAGCGGTAT 333
Db 21 GluGluAenIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40

QY 334 TCCTGGAGTCAGCTTAAAGCTGCTTCGGATACCCAGAAATATATCATGCGCTACATGATG 393
Db 41 SerTrpSerGlnLeuLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60

QY 394 GCTAAGGCACACATGATTTTCATGTTGTGAGAGAGATGATCCAGATGGACCTCATCA 453
Db 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80

QY 454 GACAGAATCTATTACCTTGCATGCTCGTGTGAGAAACAGAGAAATATACACTGTTTATCT 513
Db 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAenArgGluAenThrLeuPheTyrSer 100

QY 514 GAAATTCCTCCAAACTATCAATAGAGCAGCAGCTTAAATGCTCTCTTGGAGACCTCTTTTG 573
Db 101 GluIleProLysThrIleAsnArgAlaValLeuMetLeuSerTrpLysProLeuLeu 120

QY 574 GATCTTTTTCAGGCAACACTGGACTATGGAATGTATTCGAGAGAAAGAACTATTAGA 633
Db 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuArg 140

QY 634 GAAAGAAAACGCTATGGACAGTCGGAATTCCTTACGATTATCACCAAGGAAGTGA 693
Db 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160

QY 694 ACATTTCTGTTTCAAGCCGTAGTGAATTTATCACGTAAGAGATGGAGGCCACCAAGGA 753
Db 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180

QY 754 TTTACGCAACACCTTTAAGGCCCAATCTAGTGGAAACTAGTTGTCCTCCCAACATACGGATG 813
Db 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200

QY 814 GATCCAAATATATGCCCGCTGATCCAGACTGCTTATGCTTTTATACATAGCAAGATATT 873
Db 201 AspProLysLeuLysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220

QY 874 TGGATATCTAATCTGTACCCAGAGAGAGAGAGAGACTCATTATGTCACAAATAGCTA 933
Db 221 TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu 240

QY 934 GCCAACATGGAAGAGATGCCAGATCAGCTGAGTGCCTACCTTTGTTCTCCAAAGAGAA 993
Db 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaIleThrPheValLeuGlnGluGlu 260

QY 994 TTTGATAGATATTCTCGCTATTGTTGGTGTCCAAAGCTGAAACAACTCCAGTGGTGGT 1053
Db 261 PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly 280

QY 1054 AAAATCTTGAATTTCTATATGAAGAAAATGATGAATCTGAGGTGGAAATTTATTCATGTT 1113
Db 281 LysIleLeuArgIleLeuTyrGluGluAenAspGluSerGluValGluIleIleHisVal 300
```


QY	1114	ACATCCCTATGTTGGAACAAGAGGCGAGATTCATTCCGTTATCTCTAAACAGGTTACA	1173	QY	2194	GTGAATAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCCTAGCCTCTCTAGGT	2253
Db	301	ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr	320	Db	661	ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly	680
QY	1174	GCAATCCTAAAGTCACTTTTAAGATGTGAGAAATAATGATGTGCTGAAGGAAGATC	1233	QY	2254	TATGTGGTTGTAGTATACACAACAGGGATCCTGTACCGAGGGCTTAATATTGAAGGC	2313
Db	321	AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle	340	Db	681	TyrValValValIleAspAsnArgGlySerCysHisArgIleLysPheGluGly	700
QY	1234	ATAGATGTCATAGATAAGAACTAAATCAACCTTTTGAGATTCATTGAAGAGTTGAA	1293	QY	2314	GCCTTTAAATATAAATGGTCAAATAGAAATAGACATCAGGTGGAGGACTCCAATAT	2373
Db	341	IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu	360	Db	701	AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr	720
QY	1294	TATATGCCAGACTCGCTCAGGAGTCTCTGAGGAAAATATGCTTGGTCCATCTACTAGAT	1353	QY	2374	CTAGCTTCTCGATATGATTTCACTTTAGACTTTAGATCGTGTGGCATCCACGGTGCCTAT	2433
Db	361	TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp	380	Db	721	LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyr	740
QY	1354	CGTCCCGAGACTCGCTACAGATAGTGTGATCTCACCTGATTTATATCCAGTAGAA	1413	QY	2434	GGAGGATPACCTCTCCCTGATGGCATTAATGCAGAGGTGAGATATCTTCAGGGTTCCTATT	2493
Db	381	ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu	400	Db	741	GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle	760
QY	1414	GATGATGTTATGAAAGGAGAGACTCATTTGAGTCAAGTCCGCTGATTTCTGTGACGCCACTA	1473	QY	2494	GCTGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAACGTTATATG	2553
Db	401	AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu	420	Db	761	AlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMet	780
QY	1474	ATTATCTATGAAGAAACAACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTTTT	1533	QY	2554	GCTCACCTCACCAGATGNAACAGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAA	2613
Db	421	IleIleTyrGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe	440	Db	781	GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu	800
QY	1534	CCCCAAGTCACGAAGGAAATTTGAGTTTATTTTGGCTCTGATGCAAAACAGGTTTC	1593	QY	2614	AAGTTCCCTCTCAACCAATCGTTTACTGCTCTTACATGTTTCTCGATGAGAATGTC	2673
Db	441	ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe	460	Db	801	LysPheProSerGluProAsnArgLeuLeuLeuLeuHisGlyPheLeuAspGluAsnVal	820
QY	1594	CGTCATTTATACAAATATACATCTATTTAAAGGAAGCAATATAACGATCAGTGGT	1653	QY	2674	CATTTTGCACATACCACTATATTACTAGTATTTTGTAGTGGGCTGGAAGCCATATGAT	2733
Db	461	ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly	480	Db	821	HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp	840
QY	1654	GGGCTGCTCCTCCAGTATTTCAAGTGTCTCTATCAAGAGGAGATAGCAATACCAGT	1713	QY	2734	TTACAGATCATCTCTCAGGAGACACAGCATAAAGATTTCTGAAATCGGAGAACATTAT	2793
Db	481	GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer	500	Db	841	LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr	860
QY	1714	GGTCAATGGGAAGTCTTTGGCCGCATGGATCTAATATCCAAAGTTGATGAAGTCAGAAGG	1773	QY	2794	GAATGATCTTTTGCATCTACCTTCAAGAAACCTTGGATCAGTATTGCTGCTCTAAA	2853
Db	501	GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg	520	Db	861	GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaLeuLys	880
QY	1774	CTGGTATATTTTGAAGGCACAAAGACTCCCTTTAGAGCACTACCTCTAGTGTAGT	1833	QY	2854	GTGATA 2859	
Db	521	LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSer	540	Db	881	ValIle 882	
QY	1834	TACGTAAATCTGGAGAGTGACAAGGCTGACTGACCGTGGCTACTCATCTCTGTGTC	1893				
Db	541	TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys	560				
QY	1894	ATCAGTCAGACTGTGACTCTTTTATAGTAGTATAGTAAACCAAGAAATCCACTGT	1953				
Db	561	IleSerGlnHisCysAspPheIleSerLysTyrSerAsnGlnLysAsnProHisCys	580				
QY	1954	GTGTCCTTTACAGACTATCACTCTGAGATGACCCCAACTTGCAAAACAAAGAAATTT	2013				
Db	581	ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe	600				
QY	2014	TGGCCACCATTTTGGATTCCAGAGGTCTCTCTCTCTGACTATCTCTCCAGAAATTTTC	2073				
Db	601	TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIlePhe	620				
QY	2074	TCTTTTGAAGTACTACTGATTTTACATTTGATGGATGCTCTACAAGCCTCATGACTA	2133				
Db	621	SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu	640				
QY	2134	CAGCTCGAAAGAAATATCTCTGTGCTCTTCAATATATGGTGGTCTCAGTGCAGTTG	2193				
Db	641	GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyProGlnValGlnLeu	660				
RESULT 6							
US-10-415-122-6							
; Sequence 6, Application US/10415122							
; Publication No. US2004005369A1							
; GENERAL INFORMATION:							
; APPLICANT: THE UNIVERSITY OF SYDNEY							
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES							
; FILE REFERENCE: FFI5217							
; CURRENT APPLICATION NUMBER: US/10/415,122							
; CURRENT FILING DATE: 2003-08-07							
; NUMBER OF SEQ ID NOS: 8							
; SOFTWARE: PatentIn version 3.1							
; SEQ ID NO 6							
; LENGTH: 882							
; TYPE: PRT							
; ORGANISM: Homo sapiens							
US-10-415-122-6							
Alignment Scores:				Length:			
Pred. No.:				0			
Score:				882.00			
Percent Similarity:				100.0%			
Best Local Similarity:				100.0%			
Query Match:				85.6%			
DB:				4			
				Matches:			
				882			
				Conservative:			
				0			
				Mismatch:			
				0			
				Indels:			
				0			
				Gaps:			
				0			

US-10-825-632-2 (1-3120) x US-10-415-122-6 (1-882)

QY 214 ATGGCAGCAGCAATGGAAACAGAACAGCTGGGTGGTGTGAGATATTGAAACTGCCGACTGT 273
Db 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
QY 274 GAGGAGAAATATTGAATCACAGGATCGGCTTAAATTTGAGCCCTTTTATGTTGAGCGGTAT 333
Db 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
QY 334 TCCTGGAGTCAGCTTAAAGCTTCCTGCGATACAGAAATATCATGCTCATCATATG 393
Db 41 SerTrpSerGlnLeuLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
QY 394 GCTAAGGCACACATGATTTTCATGTTTGTGAAGAGATGATCCAGATGGACCTCATCA 453
Db 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
QY 454 GACAGATCTATTACCTTGCATGCTGTGTGAGACAGAGAAATACACTGTTTATTTCT 513
Db 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
QY 514 GAAATTCCTCCAAAATCATATGAGCAGCAGCTTTAAATGCTCTCTTGGAAAGCCTCTTTTG 573
Db 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120
QY 574 GATCTTTTTCAGGCAACACTGGACTATGGAATGTATTCTCGAGAAGAAGAACTATTAAAG 633
Db 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg 140
QY 634 GAAAGAAAACCCATTGGAAACAGTCGGAATTGCTTACGATTATACCAAGGAAGTGA 693
Db 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
QY 694 ACATTTCTGTTTCAAGCCGCTAGTGGAAATTTATACGTAAAGATGAGGGCCACACAGN 753
Db 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
QY 754 TTTACGCCAACACCTTTAGGCCCACTCTAGTGGAACTAGTCTCCACATACCGATG 813
Db 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
QY 814 GATCCAAAATTATGCCCGCTGATTCAGACTGGATGCTTTTATACATAGCAACGATATT 873
Db 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
QY 874 TGGATATCTAACATCGTAACCAAGAGAAAGGAGACTCACTTATGCAATGAGCTA 933
Db 221 TrpIleSerAsnIleValThrArgGluArgLeuThrTyrValHisAsnGluLeu 240
QY 934 GCCAACATGGAAGATGCCAGATCAGCTGAGTCGCTACCTTTGTTCTCCAAGAGAA 993
Db 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu 260
QY 994 TTTGATAGATATTCTGGCTATTGGTGGTGTCCAAAAGCTGAAACAACTCCCGAGTGGTGT 1053
Db 261 PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly 280
QY 1054 AAAATCTTAGAATTCATATGAAGAAAATGATGATCTGAGTGGAAATTTATCATGTT 1113
Db 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal 300
QY 1114 ACATCCCTTATGTTGAAACAAAGGAGGCGACATTCATTCCTGTTATCTCTAAACAGGTACA 1173
Db 301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr 320
QY 1174 GCNAATCCTTAAGTCACCTTTAAGATGTCAGAAATATGATGCTGAGAGAGGATC 1233
Db 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluArgIle 340
QY 1234 ATAGATGTCATAGAAAGCAATATTCACCTTTGAGATCTTATTTGAGAGGTGAA 1293

Db 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
QY 1294 TATATTCCAGAGCTGGATGACTCTGAGGGGAAAATATGCTTGCTCATCTACTAGAT 1353
Db 361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
QY 1354 CGCTCCACAGACTGCCCTACAGATAGTGTGTGATCTCACCTGGAATTTATATCCAGTAGAA 1413
Db 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
QY 1414 GATGATGTTATGAAAGGAGAGACTATTGAGTCAGTGCCTGTGATTTCTGTGACGCCACTA 1473
Db 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
QY 1474 ATTATCTATCAAGAAAACAACAGACATCTGGATAAATATCCATGACATCTTTCATGTTTTT 1533
Db 421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
QY 1534 CCCCAAGTCACGAAGAGAAATTTGAGTTTATTTTGGCTCTGAAATGCAAAACAGGTTTC 1593
Db 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
QY 1594 CGTCATTTATACAAAATTACATCTATTATAAGGAAAGCAAAATATAAACGATCCAGTGGT 1653
Db 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
QY 1654 GGGCTGCTCTCCCAAGTGATTTCAAGTGTCTCTATCAAGAGGAGATAGCAATTTACCAGT 1713
Db 481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer 500
QY 1714 GGTGNATGGAGATGCTTCGCGCGCATGGATCTATATCCAACTTGATGAAGTCAGAGG 1773
Db 501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520
QY 1774 CTGGTATATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACTGTAGTGTACGT 1833
Db 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSer 540
QY 1834 TAGCTAAATCTCGAGAGGTGACAAGCTGACTGACCGTGGCTACTCACATCTCTGTGTCG 1893
Db 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
QY 1894 ATCAGTCAGCACTGTGACTTCTTTTATAAGTATAGTAAACCAAGAGAATCCACACTGT 1953
Db 561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
QY 1954 GTCTCCTTTTACAAGCTATCAAGTCTCTGAAGATGACCCCACTTGCAAAAACAAAGGAATTT 2013
Db 581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
QY 2014 TGGGCCACCATTTTGGATTCAGCAGGTCCTCTTCTGACTATATCTCTCCAGAAATTTTC 2073
Db 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIlePhe 620
QY 2074 TCTTTTGAAGTACTACTGGATTTCATTTGATGGATGCTCTACAAGCCTCATGATCTA 2133
Db 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
QY 2134 CAGCCTGGAAAGAAAATATCCTACTGTGCTTTCATATATATGTTGCTCTCAGGTGCAGTTCG 2193
Db 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu 660
QY 2194 GTGAATAATCCGTTTAAAGGAGTCAGTATTTCCGCTTGAATACCTTAGCTCTTAGGT 2253
Db 661 ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly 680
QY 2254 TATGCGTGTGTAGTATAGACACAGGGATCTCTGTCCCGAGGGCTTAAATTTGAAGGC 2313
Db 681 TyrValValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly 700
QY 2314 GCCTTTAAATATAAAATCGGTCAATAGAAATTCAGATCCAGTGGAGAGCACTCCAAATAT 2373
Db 701 AlaPheLysTyrLysMetGlyGlnIleGluIleAspGlnValGluGlyLeuGlnTyr 720

QY 2374 CTAGCTCTCGATATGATTTTCATTGACTTAGTTCGTGTGGGCATCCACGGCTGCTCTAT 2433
Db 721 LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrSerTyr 740
QY 2434 GAGGATACCTCTCCCTGATGGCATTAAATGACAGAGTTCAGATATCTTCAGGGTTGCTATT 2493
Db 741 GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760
QY 2494 GCTGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGATACACGGAACTTATATG 2553
Db 761 AlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMet 780
QY 2554 GGTCACTCACCAGATGACAGGGCTATTACTTAGNCTGTGGCCATCAGACAGAA 2613
Db 781 GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu 800
QY 2614 AAGTTCCCTCTCAACCAATCGTTTACTGCTTACATGTTTCTCGATGAGAAATGTC 2673
Db 801 LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820
QY 2674 CATTTTCACATACCACTATATTAATCTAGTCTTTTGTAGTGGGCTCGAAAGCCATATGAT 2733
Db 821 HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyIleProTyrAsp 840
QY 2734 TTACAGATCTATCTCTCAGGAGACACAGACATAAGAGTTCTCGAATCGGGAGAACATTAT 2793
Db 841 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860
QY 2794 GAATGATCTTTTGCATCTACCTTCAAGAAACCTTGGATACGTTATGCTGCTCTAAAA 2853
Db 861 GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeuLys 880
QY 2854 GTGATA 2859
Db 881 ValIle 882

RESULT 7

US-10-825-632-1 ;

; Sequence 1, Application US/10825632

; Publication No. US20040191826A1

; GENERAL INFORMATION:

; APPLICANT: ABBOTT, Catherine Anne

; APPLICANT: GORRELL, Mark Douglas

; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES

; FILE REFERENCE: FCSB-100-Div. 1

; CURRENT APPLICATION NUMBER: US/10/825,632

; CURRENT FILING DATE: 2004-04-15

; PRIOR APPLICATION NUMBER: US 10/070,464

; PRIOR FILING DATE: 2002-07-18

; PRIOR APPLICATION NUMBER: PCT/AU00/01085

; PRIOR FILING DATE: 2000-09-11

; PRIOR APPLICATION NUMBER: AU P05709

; PRIOR FILING DATE: 2000-02-18

; PRIOR APPLICATION NUMBER: AU P02762

; PRIOR FILING DATE: 1999-09-10

; NUMBER OF SEQ ID NOS: 31

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1

; LENGTH: 882

; TYPE: PRT

; ORGANISM: Homo Sapiens

US-10-825-632-1

Alignment Score#:

Pred. No.: 0

Score: 882.00

Percent Similarity: 100.0%

Best Local Similarity: 100.0%

Query Match: 85.6%

DB: 4

Length: 882

Matches: 882

Conservative: 0

Mismatch: 0

Indels: 0

Caps: 0

US-10-825-632-2 (1-3120) x US-10-825-632-1 (1-882)

QY 214 ATGCACGACGAATGGAACACAGAACAGCTGGTGTGAGATATTTAAACCTCCGACTGT 273
Db 1 MetAlaAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
QY 274 GAGGAGAAATATTGAATCACAGGATCGGCTAAATTTGGAGCCCTTTTATTTGTTGAGCGTAT 333
Db 21 GluGluAsnIleGluSerGlnAspArgPheLysLeuGluProPheTyrValGluArgTyr 40
QY 334 TCCTGGAGTCAGCTTAAAAAGCTGCTGCGATACAGAAATATATCATGGCTACATGATG 393
Db 41 SerTrpSerGlnLeuLysLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
QY 394 GCTAAGGACCAACATGATTTTCATGTTTGTGAAGAGAAATGATCCAGATGGACCTCATCA 453
Db 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
QY 454 GACAGAATCTATTACCTTGCATGCTGTGTGAGAACAGAGAAATATACACTGTTTATTCT 513
Db 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
QY 514 GAAATTTCCAAAACATCAATAGACGACAGCTTTAATGCTCTCTTGAAGCCCTTTTGT 573
Db 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTyrLysProLeuLeu 120
QY 574 GATCTTTTTCAGGCAACACATGGAATGTAATCTTCGAGAGAAAGAACTATTAGA 633
Db 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg 140
QY 634 GAAAGAAAACGATTTGGAACAGCTCGAATGCTTCTTACGATATATCACCAAGGAGTGA 693
Db 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
QY 694 ACATTTCTGTTTCAAGCCGCTAGTGAATTTATTCAGTAAAGATGGAGGGCCACAGGA 753
Db 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
QY 754 TTTACGCAACAACTTTAAGGCCCAATCTAGTGGAAACTAGTGTGCCCAACATACGATG 813
Db 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
QY 814 GATCCAAAATTTATGCCCGCTGATCCAGACTGGATGCTTTTATATACATAGCAACGATAT 873
Db 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
QY 874 TGGATATCTAACATCGTAAACAGAGAAAGAGAGACTCATTATGTGCACATGAGCTA 933
Db 221 TrpIleSerAsnIleValThrArgGluGluArgArgLeuThrTyrValHisAsnGluLeu 240
QY 934 GCCAATGTGAAGAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCCAAGAA 993
Db 241 AlaAsnMetGluGluAspAlaArgSerArgIleValAlaThrPheValLeuGlnGluGlu 260
QY 994 TTTGATAGATATCTGCTGCTATTCGTGTTTCCAAAGCTGAAACAACTCCCAAGTGGTGT 1053
Db 261 PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly 280
QY 1054 AAAATTTCTTGAATTTCTATATGAAGAAATGTATGATCTGAGGTGGAATTTATTCATGTT 1113
Db 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal 300
QY 1114 ACATCCCTCTATGTTGGAACACAGAGGGGAGATTTCAATCCGTTATCTCTAAACAGGTACA 1173
Db 301 ThrSerProMetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThr 320
QY 1174 GCAATCTCTAAAGTCATTTTAAAGATGTGAGAAATATGATTTGATGCTGAGAGAGGATC 1233
Db 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
QY 1234 ATAGATGTCATAGTAAGGAACTTAATCAACCTTTTGAGATTTCTATTGAGAGGCTTGA 1293
Db 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360

QY 1294 TATATTGCCAGCTGGATCGACTCCTGAGGGAATATGCTTGCTCCATCCTACTAGAT 1353
Db TyrlleAlaargAlaGlyThrProGluGlyLysTyrrAlaThrSerIleLeuLeuAsp 380
QY 1354 CGCTCCAGACTCCCTACAGATAGTGTGATCTACCTGAATATTATATCCAGTAGAA 1413
Db ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
QY 1414 GATGATGTTATGGAAGCGCAGAGACTCATGTAGTCAGTCCCTGATCTTCGTCAACCCACTA 1473
Db AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
QY 1474 ATTATCTATGAAGAAACAAACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTTT 1533
Db IleIleTyrrGluGluThrThrAspIleThrIleAsnIleHisAspIlePheHisValPhe 440
QY 1534 CCCCAAGTCACGAAGAGGAATTTGAGTTTATTTTTGCTCTGAAATGCAAAACAGGTTTC 1593
Db ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
QY 1594 CGTCATTTATACAAATTTACATCTATTTTAAAGGAAGCAATATAACGATCCAGTGGT 1653
Db ArgHisLeuTyrrLysIleThrSerIleLeuLysGluSerLysTyrrLysArgSerSerGly 480
QY 1654 GGGCTGCTCTCCCAAGTGATTTCAAGTGTCTCTATCAAGAGGAGATAGCAATACCAGT 1713
Db GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer 500
QY 1714 GGTCAATGGGAAGTCTTGGCGGCATGGATCTAATATCCAGTTGATGAAGTCAGAGG 1773
Db GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520
QY 1774 CTGGTATATTTGAGCCACCAAGACTCCCTTTAGAGCATCACTGTAGTACGTACGT 1833
Db LeuValTyrrPheGluGlyThrLysAspSerProLeuGluHisLysLeuTyrrValValSer 540
QY 1834 TACGTAAATCCTGGAGAGTGAACAGGCTGACTGACCCGTGCTACTCACATTTCTGTGTC 1893
Db TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrrSerHisSerCysCys 560
QY 1894 ATCAGTCAGCACTGTGACTCTTTTATAAGTAAAGTATAGTAAACCGAAGAATCCACACTGT 1953
Db IleSerGlnHisCysAspPhePheIleSerLysTyrrSerAsnGlnLysAsnProHisCys 580
QY 1954 GTGTCCCTTTACAAGCTATCAAGTCTCTGAAGATGACCCCACTTGCAAAACAAAGGAATTT 2013
Db ValSerLeuTyrrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
QY 2014 TGGGCCACCATTTTGGATTCAGCAGGTCTCTTCTGACTATATCTCTCCAGAAATTTTC 2073
Db TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrrThrProGluIlePhe 620
QY 2074 TCTTTTGAAGTACTACTCGATTTACATTTGATGGATGCTCTACAGCCCTCATGATCTA 2133
Db SerPheGluSerThrThrGlyPheThrLeuTyrrGlyMetLeuTyrrLysProHisAspLeu 640
QY 2134 CAGCCTGGAAGAAATATCTACTGTGTGTCTATATATGTGTCTCTCAGGTGCAGTTG 2193
Db GlnProGlyLysLysTyrrProThrValLeuPheIleTyrrGlyLysProGlnValGlnLeu 660
QY 2194 GTGAATATCGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCTAGCTCTCTAGGT 2253
Db ValAsnAsnArgPheLysGlyValLysTyrrPheArgLeuAsnThrLeuAlaSerLeuGly 680
QY 2254 TATGTGTTGTAGTATAGAACACAGGGGATCTCTGTCAACGAGGGCTTAAATTTGAAGGC 2313
Db TyrValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly 700
QY 2314 GCCTTTAAATATATAATGGGTCAATATAGAAATTTGACGATCAGGTGGAGGACTCCAATAT 2373
Db AlaPheLysTyrrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyrr 720
QY 2374 CTAGCTTCTCGATATGATTTTCAATTGACTTAGATCGTGTGGGCATCCACGGCTGGTCTAT 2433

Db 721 LeuAlaSerArgTyrrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyrr 740
QY 2434 GGAGGATACCTCTCCCTGATGGCATTAAATGCAGAGTCAAGATATCTTCAGGGTTGCTATT 2493
Db GlyGlyTyrrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760
QY 2494 GCTGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAACGTTATATG 2553
Db AlaGlyAlaProValThrLeuTrpIlePheTyrrAspThrGlyTyrrThrGluArgTyrrMet 780
QY 2554 GGTCAACCTCTCACAGAAATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAACGACAA 2613
Db GlyHisProAspGlnAsnGluGlnGlyTyrrTyrrLeuGlySerValAlaMetGlnAlaGlu 800
QY 2614 AAGTTTCCCTCTGAAACCAATCGTTTACTCTTACATGTTTCTCGATGAGAATGTC 2673
Db LysPheProSerGluProAsnArgLeuLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820
QY 2674 CATTTTCACATACCACTGATATATTACTAGTTTTTTTAGTGAGGCTGGAAAGCCATATGAT 2733
Db HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrrAsp 840
QY 2734 TTACAGATCTATCCTCAGGAGACACACACATAAGAGTTCTCTGAATCCGGAGAACATTAT 2793
Db LeuGlnIleTyrrProGlnArgHisSerIleArgValProGluSerGlyGluHisTyrr 860
QY 2794 GAATGATCTTTTGCACTACCTTCAAGAAACCTTTGGATCACGATTGCTGCTCTCAAAA 2853
Db GluLeuHisLeuLeuHisTyrrLeuGlnGluAsnLeuGlySerArgGileAlaAlaLeuLys 880
QY 2854 GTGATA 2859
Db 881 ValIle 882
RESULT 8
US-10-982-512-1
; Sequence 1, Application US/10982512
; Publication No. US2005059081A1
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/10/982,512
; CURRENT FILING DATE: 2004-11-05
; PRIOR APPLICATION NUMBER: US/09/976,674
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 882
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-982-512-1
Alignment Scores:
Pred. No.: 0 Length: 882
Score: 882.00 Matches: 882
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 85.6% Indels: 0
DB: 5 Gaps: 0
US-10-825-632-2 (1-3120) x US-10-982-512-1 (1-882)
QY 214 ATGGCAGCAGCAATGAAACAGAACAGCTGGGTGTGTGAGATATTTGAACTGCGGACTGT 273
Db 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20

QY	274	GAGGAGAAATTGAATCACAGATCGGCCTAAATTTGGAGCGCTTTTATGTTGAGCGGTAT	333	QY	1354	CGCTCCCAAGACTCGCCTACAGATAGTGTGATCTCACCCTGAATTTATTTATCCAGTAGAA	1413
Db	21	GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr	40	Db	381	ArgSerGlnThrArgLeuGlnIleValIleuIleSerProGluLeuPheIleProValGlu	400
QY	334	TGCTGGAGTCACGCTTAAAAAGCTGCTCCCGATACACAGAAAATATCATGGCTACATGATG	393	QY	1414	GATGATGTTATGGAAGCAGAGACTCAATCAGTCACTGCCTGATTTCTGTGAGCCACTA	1473
Db	41	SerTrpSerGlnLeuLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet	60	Db	401	AspAspValMetGluArgGlnArgLeuIleGluSerValProaspSerValThrProLeu	420
QY	394	GCTAAGGCCACCATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGGACCTCATCA	453	QY	1474	ATTATCTATGAAGAAACACAGACATCTGTGATAAATATCCATGACATCTTTTCATGTTTT	1533
Db	61	AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProaspGlyProHisSer	80	Db	421	IleIleTyrGluGluThrThrAspIleThrPheAsnIleHisAspIlePheHisValPhe	440
QY	454	GACAGAAATCTATTACTCTGCCATGCTGCTGAGACAGAGAAATACACTGTTTTATTCT	513	QY	1534	CCCCAAGTCACGAAGAGAAATAGAGTTTATTTTTCCTCTGATTCACAAACAGGTTTC	1593
Db	81	AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer	100	Db	441	ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe	460
QY	514	GAATTTCCCAAACTATCAATAGACAGCAGCTCTTAATGTCTCTCTGGAGCCCTCTTTG	573	QY	1594	CGTCATTTATACAAAATTACATCTATTATAAGAAAGCAATAATAAACCATCCAGTGGT	1653
Db	101	GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu	120	Db	461	ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly	480
QY	574	GATCTTTTTCAGCAACACTGGAATGATTTCTCGAGAGAGAACTATTAAAGA	633	QY	1654	GGCTCGCTCGCTCCCAAGTGATTTCAAGTGTCTCTATCAAGAGAGAGATAGCAATTTACAGT	1713
Db	121	AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuLeuArg	140	Db	481	GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer	500
QY	634	GAAGAAGAACGATTGGACACTCGGAATTCCTTACGATTATCACCAAGCAAGTGA	693	QY	1714	GGTGAATGGGAAGTTCTTTGGCCGCATCGATCTTAATATCCAAGTTGATGAAGTCAGAGG	1773
Db	141	GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly	160	Db	501	GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg	520
QY	694	ACATTTCTGTTTCAAGCCCGTAGTGGAAATTTATCACGTAAAGATGAGGGCCACAAAG	753	QY	1774	CTGGTATATTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTCAGTAGTCAGT	1833
Db	161	ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly	180	Db	521	LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSer	540
QY	754	TTTACGCAACACTTTAAGGCCAATCTAGTGGAAACTAGTTGTCCTCCACATACGGATG	813	QY	1834	TAGCTAAATCCTGGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACATCTTCTGTCG	1893
Db	181	PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet	200	Db	541	TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys	560
QY	814	GATCCAAAATTATGCCCCCTGATCCAGACTGGATTGCTTTTATACATAGCAACGATATT	873	QY	1894	ATCAGTCAGCACTGTGACTCTTTTATAAGTAAAGTATAGTAAACCAAGAGAAATCCACACTGT	1953
Db	201	AspProLysLeuCysProAlaAspProaspTrpIleAlaPheIleHisSerAsnAspIle	220	Db	561	IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys	580
QY	874	TGGATATCTAACATCTGAACAGAGAAAGAGACTCACTTATGTGCACATAGACTA	933	QY	1954	GTGTCCCTTTTACAAAGCTATCAAGTCTCTGAGATGACCAACTGCAAAACAAAGAAATTT	2013
Db	221	TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu	240	Db	581	ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe	600
QY	934	GCCAAACATGAAGAGATGCCAGATCAGCTGGAGTCGCTACTCTTGTTCCTCCAAAGAA	993	QY	2014	TGGGCCACCACTTTTGATTCAGCAGCTCTCTCTCTGACTATATCTCTCCAGAAATTTTC	2073
Db	241	AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu	260	Db	601	TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIlePhe	620
QY	994	TTTGATAGATTTCTGGCTATTGGTGTGTCCAAAAGCTGAAACAACTCCCGTGGTGGT	1053	QY	2074	TCCTTTGAAAGTACTACTGGATTTACATTTGATGGGATGCTCTACAGCCCTCATGATCTA	2133
Db	261	PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly	280	Db	621	SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu	640
QY	1054	AAAATTTCTAGAAATCTATATGAAGAAATGATGAATCTGAGGTGGAATTTATTCATGTT	1113	QY	2134	CAGCTCGGAAGAAAATATCTCTACTGTGCTTTTCATATATGTTGGTCTCCTCAGGTCCAGTTG	2193
Db	281	LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal	300	Db	641	GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu	660
QY	1114	ACATCCCTCATTTGTGAAAACAAGGAGGCGAGATTTCATTCGCTTATCTCTAAACAGGTACA	1173	QY	2194	GTGAATAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATPACCTAGCCTCTCTAGGT	2253
Db	301	ThrSerProMetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThr	320	Db	661	ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly	680
QY	1174	GCAATCCTAAAGTCACCTTTTAAGATGTCAGAAAATGATGATGCTGAGGAAGGATC	1233	QY	2254	TATGTGTTGTAGTGTAGACAAACAGGGATCCTGTCCCGAGGCTCTAAATTTGAAGGC	2313
Db	321	AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle	340	Db	681	TyrValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly	700
QY	1234	ATGATCTCATAGTAGGAACCTAATTCACCTTTTGGATCTCTATTTTGAAGGAGTTGAA	1293	QY	2314	GCCTTTTAAATATAAATGGGTCAAAATAGAAATGAGTACGCTGAGGAGGACTCCAATAT	2373
Db	341	IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu	360	Db	701	AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr	720
QY	1294	TATATTGCCAGAGCTGGATCGACTCTCTGAGGGGAAAATATGCTTGCTCCATCTCTAGAT	1353	QY	2374	CTAGCTCTCGATATGATTTTCATTGACTTAGATCGTGTGGGCATCCACGGCTGCTCTAT	2433
Db	361	TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp	380	Db	721	LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyr	740
				QY	2434	GGAGGATACCTCTCCCTGATGGCAATTAATGCAGAGGTCAGATATCTTTCAGGGTTGCTATT	2493

```
Db 741 GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760
QY 2494 GCTGGGCCCCAGTCACTCTGTGATCTTCTATGATACAGGATACACGACGTTATATG 2553
Db 761 AlaGlyAlaProValThrLeuTyrIlePheTyrAspThrGlyThrGluArgTyrMet 780
QY 2554 GGTCACTCAGCAGATGAACAGGCTATTACTTATAGATCTGTGGCCATCCAGCAGAA 2613
Db 781 GlyHisProAspGlnGlnGlnGlyTyrLeuGlySerValAlaMetGlnAlaGlu 800
QY 2614 AAGTCCCTCTGAACCAATCGTTTACTGCTCTTACATGTTCTCTGGATGAGAATGTC 2673
Db 801 LysPheProSerGluProGlnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820
QY 2674 CATTTGCACATACCACTATATTACTGAGTCTTTTGTAGGGCTGGAAAGCCCATATGAT 2733
Db 821 HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840
QY 2734 TTACAGATCTATCTCAGGAGAGACACAGCATAGAGTTCTCGAATCGGAGAACATTAT 2793
Db 841 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860
QY 2794 GAATCGATCTTTTGCACTACCTTCAAGAAACCTTGATCACGATTTGTGCTCTAAAA 2853
Db 861 GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeuLys 880
QY 2854 GTGATA 2859
Db 881 ValIle 882

RESULT 9
US-09-976-674-19
; Sequence 19, Application US/09976674
; Patent No. US20020115843A1
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2006-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19
; LENGTH: 658
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-19

Alignment Scores:
Pred. No.: 0 Length: 658
Score: 655.00 Matches: 655
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 63.6% Indels: 0
DB: 3 Gaps: 0

US-10-825-632-2 (1-3120) x US-09-976-674-19 (1-658)
QY 214 ATGGCAGGCAATGGAACAGACAGCTGGGTGTGAGATATTTGAAACTCGGACTGT 273
Db 1 MetaAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
QY 274 GAGGAGAATATTGAATCACAGGATCGGCCCTAAATTTGAGCGCTTTTATGTTGAGCGGTAT 333
Db 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
QY 334 TCCTGGAGTCAGCTTAAAAAGCTGTTGCCGATACCCAGAAATATCATGGCTACATGATG 393
```

```
Db 41 SerTrpSerGlnLeuLysLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
QY 394 GCTAAGGCACACATGATTTTCATGTTTGTGAAGAGAAATGATCCAGATGCGACCTCATCA 453
Db 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
QY 454 GACAGAAATATTACCTTGCATGTCGTGTGAGAACAGAGAAAAATACACTGTTTTATTCT 513
Db 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
QY 514 GAAATTCCTCAAACTATCAATAGACAGCAGTCTTAATGCTCTCTTGAAGCCTCTTTTG 573
Db 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120
QY 574 GATCTTTTTCAGGCAACACTGGACTATGGAATGTATTCGAGAGAGAACTAATTAAGA 633
Db 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg 140
QY 634 GAAAGAAAACGCACTCGAACACTCGAATTCCTTACGATTATATCCAGGAGAACTGGA 693
Db 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
QY 694 ACATTTCTGTTTCAAGCCGCTAGTGAATTTATCAGCTTAAAGATGAGAGGCCCAAGGA 753
Db 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
QY 754 TTTACGCAACACCTTTAAGCCCAATCTAGTGGAACTAGTTGTCCCAACATACGGATG 813
Db 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
QY 814 GATCCAAATATTGCCCGCTGATCCAGACTGGATGCTTTTATACATAGCAACGATATT 873
Db 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
QY 874 TGGATATCTAACATCGTAAACAGAGAAAGAGAGACTCACTTATGTGCACAATGAGCTA 933
Db 221 TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu 240
QY 934 GCCAATCGAAGAAAGATGCCAGATTCAGCTGGAGTCGCTACCTTTGTTCCTCAAGAGAA 993
Db 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu 260
QY 994 TTGTATGATATTTCTGGCTATTGGTGTGTCTCAAAAGCTGAAACAACTCCCACTGGTGGT 1053
Db 261 PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly 280
QY 1054 AAAATTTTAGAATTTCTATATGAAGAAATGATGAATCTCAGGTGGAAATTTATTCATGTT 1113
Db 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHieVal 300
QY 1114 ACATCCCTTATGTTGAAACAGAGGCGCAGATTCATTCCGTTTATCTCTAAACAGGTACA 1173
Db 301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr 320
QY 1174 GCAAAATCCTAAAGTCACATTTTAAAGATGTCAGAAATATGATGATCTGAGGAGAGATC 1233
Db 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
QY 1234 ATAGATGTCATAGATAAGGAACTAATTCACACCTTTTGAGATTCATATTCAGAGAGTTGA 1293
Db 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
QY 1294 TATATTGCCAGCTGGATGGAATCTGAGGAAATATGCTTGGTCCATCTCTACTAGAT 1353
Db 361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
QY 1354 CGCTCCGAGCTCGCTACAGATAGTTGATCTCATCTGAATATTATTTATCCAGPAGAA 1413
Db 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGlnLeuPheIleProValGlu 400
QY 1414 GATGATGTTTGGAAAGCGCAGAGACTCATTTGAGTCAGTGTGATTTCTGTGACGCCACTA 1473
```

Db	401	A4pAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu	420
QY	1474	AATATCATGTAAGAAACACACAGACATCTGGATAAAATATCCATGACATCTTTTCATGTTTTT	1533
Db	421	IleIleTyrGluGluThrThrAspIleTrpIleasnIleHisAspIlePheHisValPhe	440
QY	1534	C6CCAAAGTCACGAAGAGGAAATTGAGTTTATTTTGGCTCTCTGAATGCAAAACAGGTTTC	1593
Db	441	ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe	460
QY	1594	CGTCATTATATACAAANNTACATCTATTTTAAAGGAAGCAAAATATAAAGATCCAGTGGT	1653
Db	461	ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly	480
QY	1654	GGCTGCCTGCTCCAAGTGATTTCCAAGTGCCTATCAAAGAGGAGATACCAATTACCAAGT	1713
Db	481	GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer	500
QY	1714	GGTGAATGGGAAGTTCTTGGCGGCATCGGATCTAAATATCCAAGTTGATGAAGTCAGAAGG	1773
Db	501	GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg	520
QY	1774	CTGCGTATATTTTGAAGGCACAAAGACTCCCTCTTAAAGACATCACTGTAGCTAGTCAGT	1833
Db	521	LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSer	540
QY	1834	TACGTAATCTCGGAGAGTGACAAGCTGACTGACCGTGGCTACTCATCTTCTGCTGC	1893
Db	541	TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys	560
QY	1894	ATCAGTCAGCACTGTGACTCTCTTTATAGTAAGTATAGTAACACAGAAGATCCACTGT	1953
Db	561	IleSerGlnHisCysAspPheIleSerLysTyrSerAsnGlnLysAsnProHisCys	580
QY	1954	GTGTCCTCTTACAGCTATCAAGTCCTGAAGATGACCCAACTTGCAAAACAAAGGAATTT	2013
Db	581	ValSerLeuTyrLysLysLeuSerSerProGluAspProThrCysLysThrLysGluPhe	600
QY	2014	TGGCGCACCATTTTGGATTTCAGCAGGTCCTCTTCTTGACTATATCTCTCCAGAAATTTTC	2073
Db	601	TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIlePhe	620
QY	2074	TCCTTTGAAAGTACTACTGGAATTTACATTTGATGGGATGCTCTACAAGCCTCATGATCTA	2133
Db	621	SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu	640
QY	2134	CAGCTGTGAAAGAANAATATCTTACTGTGCTGCTTTCATATATGTTGGT	2178
Db	641	GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGly	655
RESULT 10			
US-10-982-512-19			
; Sequence 19, Application US/10982512			
; Publication No. US2005005981A1			
; GENERAL INFORMATION:			
; APPLICANT: Qi, Steve			
; APPLICANT: Akinsanya, Karen			
; APPLICANT: Riviere, Pierre			
; APPLICANT: Junien, Jean-Louis			
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV			
; FILE REFERENCE: 70669			
; CURRENT APPLICATION NUMBER: US/10/982,512			
; CURRENT FILING DATE: 2004-11-05			
; PRIOR APPLICATION NUMBER: US/09/976,674			
; PRIOR FILING DATE: 2001-10-12			
; PRIOR APPLICATION NUMBER: US 60/240,117			
; PRIOR FILING DATE: 2000-10-12			
; NUMBER OF SEQ ID NOS: 61			
; SOFTWARE: PatentIn version 3.1			
; SEQ ID NO 19			
; LENGTH: 658			
; TYPE: PRT			
; ORGANISM: Homo sapiens			


```
Db 301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgGlyProLysThrGlyThr 320
QY 1174 GCAATCTTAAGTCACCTTTTAAGATGTCAGAAATATGATTTGATGCTGGAAGGATC 1233
Db 321 AlaAsnProLysValThrPheLysMetSerGluLeuMetIleAspAlaGluGlyArgIle 340
QY 1234 ATAGATGTCATAGATAAGGAACATAATCAACCTTTTGAGATTCTATTTTGAAGGAGTTGAA 1293
Db 341 IleAspValIleAspLysGluLeuIleGlnProPheGluLeuPheGluGlyValGlu 360
QY 1294 TATATTCCAGAGCTGGATCGACTCTCTGAGGGAATATGCTTGCTTCATCTACTAGAT 1353
Db 361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
QY 1354 CGCTCCAGAGCTCGCCCTACAGATAGTGTGATCTCACCTGGAATATTATCCAGTAGAA 1413
Db 381 ArgSerGlnThrArgLeuGlnIleValLeuLeuIleSerProGluLeuPheIleProValGlu 400
QY 1414 GATGATGTTATGGAAGGCAGAGACTCATTTGAGTCAGTGCCTGATTCGTGACGCCACTA 1473
Db 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
QY 1474 ATTATCTATGAAGAAACAACAGACATCTGATAAATATCCATGACATCTTTTCATGTTT 1533
Db 421 IleIleTyrGluGluThrThrAspIleTyrIleAsnIleHisAspIlePheHisValPhe 440
QY 1534 CCCCAAGTCACGAGGAAATGAGTTTATTTTGGCTCTGATGCAACCAACAGGTTTC 1593
Db 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
QY 1594 CGTCATTTATACAAATATACATCTATTTTAAGGAAGCAATATAACGATCCAGTGGT 1653
Db 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
QY 1654 GGCCTGCTGCTCCAGTGATTTCAAGTGTCTCTATCAAGAGGAGATAGCAATTTACAGT 1713
Db 481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer 500
QY 1714 GGTGAATGGGAAGTCTTGTGCGGCATGGATCTAATATCCAAAGTTGATGAAGTCAGAAG 1773
Db 501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520
QY 1774 CTGCTATATTTGAAGGCACCAAGACTCCCTCTTAGAGCATCACCTGTACGTAGTCAGT 1833
Db 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSer 540
QY 1834 TACGTAATCTCGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACATTTCTTGTCG 1893
Db 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
QY 1894 ATCAGTCAGCACTGTGACTTCTTTATAAGTAAAGTATAGTAACCAAGAAATCCACACTGT 1953
Db 561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
QY 1954 GTGTCCTTTTACAGCTATCAAGTCCTGAAGATGCCCACTTGCACAAACAAAGGAATTT 2013
Db 581 ValSerLeuTyrLysLeuSerSerProGluLysAspAspProThrCysLysThrLysGluPhe 600
QY 2014 TGCGCCACCAATTTGGATTTCAGAGTCCTCTCTGACTATCTACTCTCCAGAAATTTTC 2073
Db 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIlePhe 620
QY 2074 TCTTTTGAAGTACTACTGATTTACATTTGATGGATGCTCTACAGCCCTCATGATCTA 2133
Db 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
QY 2134 CAGCTCGGAAGAAATATCTACTGTGCTGTTTCATATATGGTGGT 2178
Db 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGly 655
```

RESULT 11
US-09-976-674-11

```
; Sequence 11, Application US/09976674  
; Patent No. US20020115943A1  
; GENERAL INFORMATION:  
; APPLICANT: Q1, Steve  
; APPLICANT: Akinsanya, Karen  
; APPLICANT: Riviere, Pierre  
; APPLICANT: Junien, Jean-Louis  
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV  
; FILE REFERENCE: 70669  
; CURRENT APPLICATION NUMBER: US/09/976,674  
; CURRENT FILING DATE: 2001-10-12  
; PRIOR APPLICATION NUMBER: US 60/240,117  
; PRIOR FILING DATE: 2000-10-12  
; NUMBER OF SEQ ID NOS: 61  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 11  
; LENGTH: 661  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; US-09-976-674-11
```

```
Alignment Scores:  
Pred. No.: 0 Length: 661  
Score: 655.00 Matches: 655  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 63.6% Indels: 0  
DB: 3 Gaps: 0
```

US-10-825-632-2 (1-3120) x US-09-976-674-11 (1-661)

```
QY 214 ATGCACAGCAATGGNAACAGACGCTGGTGTGAGATATTTGAACTCCGACTGT 273  
Db 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20  
QY 274 GAGAGAAATATGAAATCACAGATCGGCCTAAATTGGAGCCCTTTTATGTTGAGCGGTAT 333  
Db 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40  
QY 334 TCCTGGAGTCAGCTTAAAAAGCTGCTGCCGATACACAGAAAAATATCATGGCTACATGATG 393  
Db 41 SerTrpSerGlnLeuLysLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60  
QY 394 GCTAAGGCACCAATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGGACCTCATTTCA 453  
Db 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80  
QY 454 GACAGAAATCTATTACCTTGCATGCTGCTGGTGAGAACACAGAGAAATACACTGTTTATTCT 513  
Db 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100  
QY 514 GAAATTCCTCAAACTATCAATAGACAGCAGTCTTAATGCTCTCTTGAAGCCTCTTTTG 573  
Db 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120  
QY 574 GATCTTTTTCAGGCACACCTGGACTATGGAATGATTTCTTCGAGAGAGAACTATTAAAGA 633  
Db 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuLeuArg 140  
QY 634 GAAAGAAACGCATTCGACAGTCGGAATTCCTTACGATTATCACCAGGAGGTGGA 693  
Db 141 GluArgArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160  
QY 694 ACATTTCTGTTTCAGCCGGTAGTCGAATTTATCAGTAAAGATGAGGAGGCCACAGGA 753  
Db 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180  
QY 754 TTTTACGCAACACCTTTAAGGCCCAATCTAGTGGAAACTAGTGTGCCCAACATACGATG 813  
Db 181 PheThrGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200  
QY 814 GATCCAAAAATTATGCCCCGCTGATCCAGACTGGATGCTTTTATACATAGCAACGATTT 873
```


Db	201	AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle	220
QY	874	TGGATATCTAAACATCGTAAACAGAGAAGAGAGACTCACTTATGTGCGACAAATGAGCTA	933
Db	221	TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu	240
QY	934	GCCAACTGGAAGAGATGCCAGATCAGCTGAGTGGTACCTCTTGTCTCCAGAGAA	993
Db	241	AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu	260
QY	994	TTTGATAGATATTCTGGCTATTGGTGTGTCAAAAGCTGAAACAACTCCCGAGTGGT	1053
Db	261	PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly	280
QY	1054	AAAATCTTAGAATCTATAGAGAAAATGATGAATCTGAGGTGGAAATTTATTCATGTT	1113
Db	281	LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal	300
QY	1114	ACATCCCTATGTTGGAACAGAGGCGAGATTCATTCGGTTATCTCTAAACAGGTACA	1173
Db	301	ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr	320
QY	1174	GCAATCTTAAAGTCACCTTTTAAGATGTCAGAAATAATGATTGATGCTGAAGAGAGATC	1233
Db	321	AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle	340
QY	1234	ATAGATGTCTAGATAAGAACTAATTCACCTTTTGGATTTCTATTGGAAGGAGTTGAA	1293
Db	341	IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu	360
QY	1294	TATATTCGACAGCTGGATGACTCCTGAGGAAAATATGCTTGGTCCATCTCTACTAGAT	1353
Db	361	TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp	380
QY	1354	CGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGMAATTTATTTATCCAGTAGAA	1413
Db	381	ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu	400
QY	1414	GATGATGTTATGGAAGGCGAGACTCATTGAGTGCCTGATCTGTGACGCCACTA	1473
Db	401	AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu	420
QY	1474	ATTATCTATGAAGAAACACAGACATCTGGATAAAATATCCATGACATCTTTTCATGTTTT	1533
Db	421	IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe	440
QY	1534	CCCAAAGTCACGAAGAGAAATTTGAGTTTATTTTGGCTCTGAAATGCAAAACAGGTTTC	1593
Db	441	ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe	460
QY	1594	CGTCATTTATACAAAATTACATCTATTTTAAAGGAAGCAATATATAACGATCCAGTGGT	1653
Db	461	ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly	480
QY	1654	GdGCTGCTGCTCCAAAGTATTTCAAGTGTCTTATCAAGAGGAGATAGCAATTTACAGT	1713
Db	481	GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer	500
QY	1714	GdTGAAATGGAAAGTCTTGGCCGGCATGGATCTAATATCCAAGTTGATCAAGTCAGAGG	1773
Db	501	GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg	520
QY	1774	CTGGTATATTTTGAAGGACCAAGACTCCCTTTTAGAGCATCACCTGTACGTACTGACT	1833
Db	521	LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSer	540
QY	1834	TACGTAATCTCGAGAGGTGACAAAGCTGACTGACCGTGGCTACTCAATTTCTGTGTC	1893
Db	541	TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys	560
QY	1894	ATCAGTCAGCATGTGACTCTCTTTATAAGTAAAGTATAGTAACAGAGAATCCACACTGT	1953
Db	561	IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys	580
QY	1954	CTGTCCCTTTTACAAGCTATCAAGTCTCTGAAGTACCCAACTTGCAAAACAAAGAAATTT	2013
Db	581	ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe	600
QY	2014	TGGGCCACCAATTTTGGATTTCAGAGGTCTCTTCTCTGACTATCTCTCCAGAAATTTTC	2073
Db	601	TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIlePhe	620
QY	2074	TCCTTTGAAAGTACTACTGGATTTACATTGTATGGATGCTCTCAACGCCTCATGATCTA	2133
Db	621	SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu	640
QY	2134	CAGCTCGAAAGAAATATCTACTGTGCTGTTTCATATATGTTGGT	2178
Db	641	GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGly	655
RESULT 12			
US-10-982-512-11			
; Sequence 11, Application US/10982512			
; Publication No. US20050059081A1			
; GENERAL INFORMATION:			
; APPLICANT: Qi, Steve			
; APPLICANT: Akinsanya, Karen			
; APPLICANT: Riviere, Pierre			
; APPLICANT: Junien, Jean-Louis			
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV			
; FILE REFERENCE: 70669			
; CURRENT APPLICATION NUMBER: US/10/982,512			
; CURRENT FILING DATE: 2004-11-05			
; PRIOR APPLICATION NUMBER: US/09/976,674			
; PRIOR FILING DATE: 2001-10-12			
; PRIOR APPLICATION NUMBER: US 60/240,117			
; PRIOR FILING DATE: 2000-10-12			
; NUMBER OF SEQ ID NOS: 61			
; SOFTWARE: Patent in version 3.1			
; SEQ ID NO 11			
; LENGTH: 661			
; TYPE: PRT			
; ORGANISM: Homo sapiens			
US-10-982-512-11			
Alignment Scores:			
Pred. No.:	0	Length:	661
Score:	655.00	Matches:	655
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	63.6%	Indels:	0
DB:	5	Gaps:	0
US-10-825-632-2 (1-3120) x US-10-982-512-11 (1-661)			
QY	214	ATGCAGCAGCAATGGAAACAGAACAGCTGGGTGTTGAGATATTTGAAACGCGACTGT	273
Db	1	MetAlaAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys	20
QY	274	GAGGAGAATATTGAATCACAGGATCGGCCTAAATTTGAGCCCTTTTATTTGAGCGGTAT	333
Db	21	GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr	40
QY	334	TCCTGGAGTCAGCTTAAAGAGCTGTCGCGATACAGAAAATATCATGGCTACATGATG	393
Db	41	SerTrpSerGlnLeuLysLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet	60
QY	394	GCTAAGGCACCATGATTTTCATGTTTGGAGAGAATGATCCAGATGGACCTCATCA	453
Db	61	AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer	80
QY	454	GACAGAATCTATTACCTTGGCATGCTGCTGGTGAGAACAGAGAAAATACACTGTTTATTTCT	513
Db	81	AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer	100
QY	514	GAAATTCCTCAAAACATATCAATAGACGACGAGTCTTAATGCTCTCTTTGGAAGCCTCTTTG	573

101 GluileProLysThrIleAsnArgAlaIaValLeuMetLeuSerTrpLysProLeuLeu 120
120
574 GATCTTTTTCAGGCAACACTGGACTATGGAATGATTCTCGAGAGAGACTATTAAAGA 633
121 AspLeuPheGlnAlaThrLeuAspTrpGlyMetCysArgGluGluGluLeuLeuArg 140
634 GAAAGAAAACGCATTGGAACAGTCGGAATTCCTTACGATATTCACCAAGGAAGTGA 693
141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
694 ACATTTCTGTTTCAAGCCGCTAGTGGAAATTTATACGTAAAGATGGAGGCCCAAGGA 753
161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
754 TTTACGCAACAACCTTTAAGGCCCAATCTAGTGGAACTAGTGTGTCACCAATACGGATG 813
181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
814 GATCCAAAATATTGCCCCGCTGATCCAGACTGGATTGCTTTTATACATAGCAACGATATT 873
201 AspProLysLeuCysProAlaAspProAspTrrpIleAlaPheIleHisSerAsnAspIle 220
874 TGGATATCTAACATCGTAACACAGAGAAAGAGACTCACTTATGTGCACAATGAGCTA 933
221 TrpIleSerAsnIleValThrArgGluGluArgAlaGlyValAlaThrPheValLeuGlnGlu 240
934 GCCAAACATGGAAGAAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAAGAGAA 993
241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu 260
994 TTTGATAGATATTCTGGCTATTGTTGGTCTCCAAAGCTGAAACAACTCCCAAGTGGGT 1053
261 PheAspArgTyrSerGlyTyrTrpCysProLysAlaGluThrThrProSerGlyGly 280
1054 ARAATCTTAGAATCTATAGNAAGAAATGATCTGAGGTGGAAATATTATCATGTT 1113
281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleHisVal 300
1114 ACATCCCTATGTTGAAACACAGAGGCGAGATTCATTCGTTATCTCTAAACAGGTPACA 1173
301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr 320
1174 GCAATCTTAAAGTCACTTTTAAAGATGTCAGAAATAATGATTGATCTGGAAGGAGATC 1233
321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
1234 ATAGATGTCATAGATAAGAACTAATTCACCTTTTGAGATTCCTATTGGAAGGAGTTGAA 1293
341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
1294 TATATTGCCAGAGCTGGAGTCTCTCGAGGAAAATATGCTTGGTCCATCTACTAGAT 1353
361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
1354 CGTCCCAAGTCCGCTCAGATAGTGTGATCTCACCTGAATTTATTTATCCCAAGTAGAA 1413
381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
1414 GATGATGTTATGAAAGGAGAGACTCAATTCAGTCAGTCCTGATCTGTGAGCGCCACTA 1473
401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
1474 ATTATCTATGAGAACACACACATCTGGATAATATCCATGACATCTTTCATCTTTT 1533
421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
1534 CCCCAAGTCAGAGAGGAAATTCAGTTTATTTTGGCTCTGATGCAAAACAGTTTC 1593
441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
1594 CGTCATTTATACAAATTCATCTATTATTAAGGAAAGCAAAATATAAACGATCCAGTGT 1653

461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
1654 GGGCTGCTGCTCCAGTGTCAAGTGTCAAGTGTCTTATCAAGAGAGAGATAGCAATACCAGT 1713
481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer 500
1714 GGTGAATGGGAAGTTCTTGGCCGGCATGGATCTAATATATCAAGTTGATGAAGTCAGAAG 1773
501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520
1774 CTGTATATATTTGAAGGACCAAGACTCCCTTTTAGAGCATCACTCTAGTGTAGTCACT 1833
521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSer 540
1834 TACATAAATCCTCGAGAGGTGACAGGCTGACTGACCGTGGCTACTCTACATTTCTTGCTGC 1893
541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
1894 ATCAGTCAGCACTGTGACTTCTTTATAGTAAGTATAGTAACCAAGAAATCCACACTGT 1953
561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
1954 GTGTCCCTTTTACAAGCTATCAAGTCTCTGAAGTACCACTTCGCAAAACAAGAAATTT 2013
581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
2014 TGGCCACCACTTTTGGATTTCAGAGTCTCTCTCTGACTACTCTCTCGAAATTTTC 2073
601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe 620
2074 TCTTTTGAAGTACTACTCGATTTACATTTGATGATGATGCTCTACAGCCTCATGATCTA 2133
621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
2134 CAGCTGGAAAGAAATATCCTACTGTGCTGTTCATATATGTTGGT 2178
641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGly 655

RESULT 13
US-09-976-674-7
; Sequence 7, Application US/09976674
; Patent No. US20020115843A1
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCES: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 690
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-7

Alignment Scores:
Pred. No.: 0 Length: 690
Score: 655.00 Matches: 655
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 63.6% Indels: 0
DB: 3 Gaps: 0

US-10-825-632-2 (1-3120) x US-09-976-674-7 (1-690)

Qy 214 ATGGCAGCAGCAATCGAAACAGACAGCTGGTGTTCAGATATTTGAACTCGCGACTGT 273

Db 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
Qy 274 GAGAGAAATTGAATCAGAGATCGGCTAAATTGGAGCTTTTATGTTGAGCGGTAT 333
Db 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
Qy 334 TCCTGGAGTCAGCTTAAAGAGCTGCTGCCGATACCAGAAAATATCATGGCTACATGATG 393
Db 41 SerTrpSerGlnLeuLysLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
Qy 394 GCTAAGCCACCATGATTTTCATGTTGTGAAGAGGAATGATCCAGATGACCTCATTTCA 453
Db 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
Qy 454 GACAGATCTTATACCTTCCCATGCTGCTGAGAACAGAGAAAATACACTGTTTATTTCT 513
Db 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
Qy 514 GAAATTCCTCAAACTATCAATAGACGACAGTCTTAATGCTCTCTCGAAGCCTCTTTTG 573
Db 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120
Qy 574 GATCTTTTTCAGGCAACACTGGACTATGGAATGATTTCTCAGAGAAGAAGAACTTAAAG 633
Db 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuLeuArg 140
Qy 634 GAAAGAAAGCATTGGAACAGTCGGAATTGCTTACGATTTATCAAGTAAAGATGGAGGCCACAAGGA 753
Db 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
Qy 694 ACATTTCTGTTTCAAGCCGCTAGTGGAAATTTATCAAGTAAAGATGGAGGCCACAAGGA 753
Db 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
Qy 754 TTTACGCAACAACTTTAAGGCCAATCTAGTGGAAACTAGTTGTCCTCCACATACGGATG 813
Db 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
Qy 814 GATCCAAATTAATGCCCCCTGATCCAGACTGGATGCTTTTATACATAGCAACGATATT 873
Db 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
Qy 874 TGGATATCTTAACATCGTAACAGAGAAAGAGACTCACTTATGTGCAATAGACTA 933
Db 221 TrpIleSerAsnIleValThrArgGluArgGluLeuThrTyrValHisAsnGluLeu 240
Qy 934 GCCAACATGAAGAAGATGCCAGATCAGCTGAGTCGCTACCTTTGTTCTCCAGAGAGAA 993
Db 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu 260
Qy 994 TTTGATAGATATTCTGGCTATTGCTGTCCTCAAAAGCTGAAACAACTCCCGAGTGGT 1053
Db 261 PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly 280
Qy 1054 AAAATTTCTAGAATTTCTATTAAGAAAATGATGAACTCGAGGTGGAAATTTATCATGTT 1113
Db 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal 300
Qy 1114 ACATCCCTATGTTGGAAAACAAGAGGCGAGATTCATCCGTTATCTCTAAACAGGTACA 1173
Db 301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr 320
Qy 1174 GCAATCTCAAGTCATTTTAAGATGTCAGAAAATAATGATTGATGCTGGAAGGAGATC 1233
Db 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
Qy 1234 ATAGATGTCATAGATAAGAACTAATTCACCTTTTTCAGATTTCTATTGAAAGGAGTTCAA 1293
Db 341 IleAspValIleAspLysGluLeuGlnProPheGluIleLeuPheGluGlyValGlu 360
Qy 1294 TATATTGCCAGAGCTGATGACTCTCTGAGGGAATAATGCTGCTCCATCTCTACTAGAT 1353
Db 361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380

Qy 1354 CGCTCCAGACCTCGCTACAGATAGTGTGATCTCACCTGAATTTATTTATCCAGTAGAA 1413
Db 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
Qy 1414 GATGATGTTATGAAAGGAGAGACTCATTTAGTCAGTCGCTGATTTCTGTGAGCCACTA 1473
Db 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
Qy 1474 ATTATCTATGAAGAAACAACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTTT 1533
Db 421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
Qy 1534 CCCCAGAGTCACGAAGAGAAATTCAGTTTATTTTGGCTCTGAAATGCAAAACAGGTTTC 1593
Db 441 ProGlnSerHisGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
Qy 1594 CGTCATTTATACAAAATTCATCTATTTTAAAGGAAGCAATAATAACCATCCAGTGGT 1653
Db 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
Qy 1654 GGGCTGCTCTCCAAAGTCAATTTCAAGTGTCTCTATCAAGAGAGATAGCAATTTACAGT 1713
Db 481 GlyLeuProAlaProSerAspPheLysCysProLysGluGluIleAlaIleThrSer 500
Qy 1714 GGTCAATGGGAAGTCTTGGCCGATGGATCTAATATCTCAAGTTGATGAAGTCAGAGG 1773
Db 501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArg 520
Qy 1774 CTGGTATATTTTGAAGGCCAACAGACTCCCTTTTAGAGCATCACCTGTAGTGTAGTCAGT 1833
Db 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSer 540
Qy 1834 TACGTAATCTCGAGAGGTGACAAGGCTGACTGACCGCTGCTACTCACATTTCTGCTGC 1893
Db 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
Qy 1894 ATCAGTCAGCACTGTGACTTCTTTATAGTAAGTATAGTAACCAAGAAATCCACACTGT 1953
Db 561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
Qy 1954 GTGTCCTTTTACAAGCTATCAAGTCTCTGAAGATGACCAACTTGCACAAACAAAGGAATTT 2013
Db 581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
Qy 2014 TGGGCCACCAATTTTGGATTCAGCAGGTCTCTTCTCTGACTATCTCTCCAGAAATTTTC 2073
Db 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIlePhe 620
Qy 2074 TCTTTTGAAGTACTACTGGATTTTACATTTGATGGATGCTCTACAGCCTCATGATCTA 2133
Db 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
Qy 2134 CAGCCTCGAAGAAATATCTACTGCTGCTTTCATATATGTTGGT 2178
Db 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGly 655

RESULT 14

US-10-982-512-7
; Sequence 7, Application US/10982512
; Publication No. US20050059081A1
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/10/982,512
; CURRENT FILING DATE: 2004-11-05
; PRIOR APPLICATION NUMBER: US/09/976,674
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117

Qy	2134	CAGCCTGGAAGAAATATCTACTGTGCTGTTTCATATATGTTGGT	2178
Db	641	GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGly	655

RESULT 15

```

US-09-976-674-21
; Sequence 21, Application US/09976674
; Patent No. US20020115843A1
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 21
; LENGTH: 613
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-21

```

Alignment Scores:		
Pred. No.:	0	613
Score:	607.00	Length:
Percent Similarity:	100.0%	Matches:
Best Local Similarity:	100.0%	Conservative:
Query Match:	58.9%	Mismatches:
DB:	3	Indels:
	3	Gaps:
	0	

US-10-825-632-2 (1-3120) x US-09-976-674-21 (1-613)

QY	214	ATGGCAGCAGCAATGAAACAGAACAGACGTGGGTGTGAGATATTGAAATCGCGGACTGT	273
Db	1	MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys	20
QY	274	GAGGAGAATTGAATTCAGAGATCGGCCTAAATGGAGCCTTTTATGTTGAGCCGGTAT	333
Db	21	GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyValGluAArgTyr	40
QY	334	TCTCGAGTCAGCTTAAAGAAGCTGCTCCGCATACCCAGAAAAATATCATGCTACATGATG	393
Db	41	SerTrpSerGlnLeuLysLeuLeuAlaaspThrArgLysTyrHisGlyTyrMetMet	60
QY	394	GCTAAGGCCACCATGATTTCATGTTTCTGAAGAGGAATGATCCAGATGACCTCATTTCA	453
Db	61	AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProaspGlyProHisSer	80
QY	454	GACAGAATCTATTACTTGGCCATGCTCGGTGAGAACAGAGAAATACACTGTTTTTATCT	513
Db	81	AspArgIleTyrTyLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTySer	100
QY	514	GAATTCCCAAACTATCAATAGACAGCAGCTCTTAATGCTCTCTGGAGACCTCTTTTG	573
Db	101	GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu	120
QY	574	GATCTTTTTCAGCCCAACTGCAGCATGGAATGTATTCTCGAGAAGAAGAACTATTAAAG	633
Db	121	AppLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuAArg	140
QY	634	GAAGAAGAAACGCAATTGGAAACAGTCGGGAATGCTCTTACGATTATCAACGAAGTGA	693
Db	141	GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly	160
QY	694	ACATTTCTGTTTCAAGCCGGTAGTGGAAATTTTATCACGTAAAGAATGAGGGCCACAAAG	753
Db	161	ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly	180

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model
Run on: May 2, 2006, 04:46:40 ; Search time 11.5 Seconds
(without alignments)
2467.668 Million cell updates/sec

Title: US-10-825-632-2
Perfect score: 1030
Sequence: 1 aagtgctaaagctccgagg.....agaattactataaaaaaaa 3120

Scoring table:

oligo	
Xgapop 60.0	Xgapext 60.0
Ygapop 60.0	Ygapext 60.0
Fgapop 60.0	Fgapext 7.0
Delop 60.0	Delext 60.0

Searched: 232119 seqs, 45477862 residues

Word size: 1

Total number of hits satisfying chosen parameters: 424616

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame_n2p.model -DEV=xlp
-Q=/abss/ABSSWEB_spool/US10825632/runat_01052006_110250_4729/app_query.fasta_1
-DB=Published Applications AA New -QFMT=fastan -SUFFIX=oligo.n2p.rapbn
-MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR_SCORE=quality -THR_MIN=1
-ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -HOST=abs02p
-USER=US10825632 @CGN 1 1 34 @runat 01052006 110250 4729 -NCPU=6 -ICPU=3
-NO_WMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=60 -FGAPEXT=7
-YGAPOP=60 -YGAPEXT=60 -DELOP=60 -DELEXT=60

Database : Published Applications AA New:

- 1: /SID55/ptodata/2/pubpaa/US08_NEW_PUB.pep1.*
- 2: /SID55/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 3: /SID55/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 4: /SID55/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 5: /SID55/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 6: /SID55/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 7: /SID55/ptodata/2/pubpaa/US09_NEW_PUB.pep1.*
- 8: /SID55/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 9: /SID55/ptodata/2/pubpaa/US10_NEW_PUB.pep1.*
- 10: /SID55/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
- 11: /SID55/ptodata/2/pubpaa/US11_NEW_PUB.pep1.*
- 12: /SID55/ptodata/2/pubpaa/US60_NEW_PUB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	882	85.6	882	11	US-11-151-601-20
2	109	10.6	109	11	Sequence 20, Appl
3	24	2.3	24	11	Sequence 10, Appl
4	11	1.1	99	11	Sequence 16, Appl
5	11	1.1	349	11	Sequence 11, Appl
					Sequence 1591, Ap

6	11	1.1	497	11	US-11-264-096-1593	Sequence 1593, Ap
7	11	1.1	497	11	US-11-264-096-1594	Sequence 1594, Ap
8	9	0.9	353	11	US-11-172-740-719	Sequence 719, App
9	8	0.8	49	11	US-11-071-259-3	Sequence 3, Appli
10	8	0.8	90	11	US-11-096-568A-9319	Sequence 9319, Ap
11	8	0.8	108	9	US-10-504-588-6	Sequence 6, Appli
12	8	0.8	108	11	US-11-073-420-6	Sequence 6, Appli
13	8	0.8	123	11	US-11-194-246-399	Sequence 399, App
14	8	0.8	129	11	US-11-096-568A-14806	Sequence 14806, A
15	8	0.8	129	11	US-11-096-568A-26015	Sequence 26015, A
16	8	0.8	137	11	US-11-087-099-943	Sequence 943, App
17	8	0.8	147	11	US-11-087-099-10322	Sequence 10322, A
18	8	0.8	149	11	US-11-096-568A-26340	Sequence 26340, A
19	8	0.8	151	11	US-11-096-568A-24024	Sequence 24024, A
20	8	0.8	157	11	US-11-096-568A-19057	Sequence 19057, A
21	8	0.8	164	11	US-11-096-568A-14805	Sequence 14805, A
22	8	0.8	164	11	US-11-096-568A-26014	Sequence 26014, A
23	8	0.8	174	11	US-11-072-512-2398	Sequence 2398, Ap
24	8	0.8	177	11	US-11-096-568A-14804	Sequence 14804, A
25	8	0.8	177	11	US-11-096-568A-26013	Sequence 26013, A
26	8	0.8	187	11	US-11-096-568A-24022	Sequence 24022, A
27	8	0.8	200	11	US-11-096-568A-19056	Sequence 19056, A
28	8	0.8	200	11	US-11-096-568A-26339	Sequence 26339, A
29	8	0.8	224	11	US-11-096-568A-15343	Sequence 15343, A
30	8	0.8	233	11	US-11-188-298-2947	Sequence 2947, Ap
31	8	0.8	245	11	US-11-096-568A-10277	Sequence 10277, A
32	8	0.8	246	11	US-11-096-568A-19055	Sequence 19055, A
33	8	0.8	249	11	US-11-096-568A-13087	Sequence 13087, A
34	8	0.8	266	11	US-11-096-568A-19487	Sequence 19487, A
35	8	0.8	270	11	US-11-098-686-10898	Sequence 10898, A
36	8	0.8	281	11	US-11-096-568A-15342	Sequence 15342, A
37	8	0.8	287	11	US-11-080-991-66	Sequence 66, Appl
38	8	0.8	295	11	US-11-096-568A-13086	Sequence 13086, A
39	8	0.8	303	9	US-10-517939-214	Sequence 214, App
40	8	0.8	309	11	US-11-172-740-1121	Sequence 1121, Ap
41	8	0.8	311	11	US-11-188-298-21052	Sequence 21052, A
42	8	0.8	315	11	US-11-096-568A-19486	Sequence 19486, A
43	8	0.8	317	11	US-11-184-298-19438	Sequence 19438, A
44	8	0.8	319	9	US-10-055-977-290	Sequence 290, App
45	8	0.8	323	11	US-11-096-568A-13085	Sequence 13085, A

ALIGNMENTS

RESULT 1

- US-11-151-601-20
- Sequence 20, Application US/1151601
- Publication No: US20060003413A1
- GENERAL INFORMATION:
- APPLICANT: Millennium Pharmaceuticals, Inc.
- APPLICANT: Meyers, Rachel E.
- APPLICANT: Olandt, Peter J.
- APPLICANT: Kapeller-Libermann, Rosana
- APPLICANT: Curtis, Rory A. J.
- APPLICANT: Williamson, Mark
- APPLICANT: Weich, Nadine
- TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASE, PHOSPHATASE.
- TITLE OF INVENTION: AND PROTEASE FAMILY MEMBERS AND USES THEREOF
- FILE REFERENCE: MP100-054P1RCPIOMNIDVIM
- CURRENT APPLICATION NUMBER: US/11/151.601
- CURRENT FILING DATE: 2005-06-13
- PRIOR APPLICATION NUMBER: US 10/170,789
- PRIOR FILING DATE: 2002-06-13
- PRIOR APPLICATION NUMBER: US 09/797,039
- PRIOR FILING DATE: 2001-02-28
- PRIOR APPLICATION NUMBER: PCT/US01/06525
- PRIOR FILING DATE: 2001-02-28
- PRIOR APPLICATION NUMBER: US 60/186,061
- PRIOR FILING DATE: 2000-02-29
- PRIOR APPLICATION NUMBER: US 09/882,166
- PRIOR FILING DATE: 2001-06-15
- PRIOR APPLICATION NUMBER: PCT/US01/19269
- PRIOR FILING DATE: 2001-06-15

```
; PRIOR APPLICATION NUMBER: US 60/212,078
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 09/934,406
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: PCT/US01/26052
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/226,740
; PRIOR FILING DATE: 2000-08-21
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 882
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-151-601-20

Alignment Scores:
Pred. No.: 0          Length: 882
Score: 882.00        Matches: 882
Percent Similarity: 100.0%      Conservative: 0
Best Local Similarity: 100.0%    Mismatches: 0
Query Match: 85.6%              Indels: 0
DB: 11                         Gaps: 0

US-10-825-632-2 (1-3120) x US-11-151-601-20 (1-882)
QY 214 ATGCGACGCAATGGAACAGACAGCTGGGTGTGAGATATTGAACTCGGACTGT 273
DB 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaSerPys 20
QY 274 GAGAGAAATATTGAATCACAGATCGGCTAAATTGGAGCCCTTTTATGTTGAGCGGTAT 333
DB 21 GluGluAsnIleGluSerGlnAepA-gProLysLeuGluProPheTyrValGluArgTyr 40
QY 334 TCCTGGAGTCAGCTTAAAAAGCTGCTGCCGATACCAGAAAATATCATGGCTACATGATG 393
DB 41 SerTrpSerGlnLeuLysLysLeuLeuAlaAepThrArgLysTyrHisGlyTyrMetMet 60
QY 394 GCTAAGCACCACATGATTTTCATGTTTGTGAAGAGAAATGATCCAGATGGACCTCATTTCA 453
DB 61 AlaLysAlaProHisAepPheMetPheValLysArgAsnAepProAepGlyProHisSer 80
QY 454 GACAGATCTATTACCTTGGCCATGCTGTGGTGAAGACAGAGAAAATACACTGTTTATTCT 513
DB 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
QY 514 GAAATTCCTCAAAATCATCAATAGACGACGACTATGGAATGTATCTCTTGGAAAGCCTTTTG 573
DB 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120
QY 574 GATCTTTTCAGGCAACACTGGACTATGGAATGTATCTCCGAGAGAGAACTATTAAAGA 633
DB 121 AspLeuPheGlnAlaThrLeuAepTyrGlyMetTyrSerArgGluGluLeuLeuArg 140
QY 634 GAAAGAAACGCATTGGAACAGTCGGAATTTGCTTACGATTATCACAGGAAGTGA 693
DB 141 GluArgLysArgGlyIleGlyThrValGlyIleAlaSerTyrAepTyrHisGlnGlySerGly 160
QY 694 ACATTTCTGTTTCAGCCGCTAGTGGAAATTTATCATCGTAAAGATGGAGGCCCAAGGA 753
DB 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAepGlyGlyProGlnGly 180
QY 754 TTTACGCAACACCTTTTAAAGCCCAATCTAGTGGAAACTAGTTGTGCCCAACATACGGATG 813
DB 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
QY 814 GATCCAAAATTTATGCCCGCTGATCCAGACTGGATTTGCTTTTATATCATAGCAACGATATT 873
DB 201 AspProLysLeuCysProAlaAepProAepTrpIleAlaPheIleHisSerAsnAepIle 220
QY 874 TGGATATCTTAACATCGTAACAGAGAAAGAGAGACTCATTTATGTGCCAAATGAGCTA 933
```

```
DB 221 TrpIleSerAsnIleValThrArgGluGluArgArgLeuThrTyrValHisAsnGluLeu 240
QY 934 GCCAACATGGAAGAAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAGAAGAA 993
DB 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu 260
QY 994 TTTGATAGATATTCTGGCTATTGCTGTCTCCAAAGCTGAAACCAACCTCCAGTGGTGGT 1053
DB 261 PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly 280
QY 1054 AAAATTCTTACAATTTATATGAAGAAAATGATGAATCTGAGGTGGAATATTATCATGTT 1113
DB 281 LysIleLeuArgIleLeuTyrGluGluAsnAepGluSerGluValGluIleIleHisVal 300
QY 1114 ACATCCCTCATGTTGAAACACAGAGGCGAGATTTCATTCGTTATCTCTAAACAGGTACA 1173
DB 301 ThrSerProMetLeuGluThrArgArgAlaAepSerPheArgTyrProLysThrGlyThr 320
QY 1174 GCAAAATCCTAAAGTCACCTTTTAAGATGTCAGAAAATTAATGATTGATGCTGGAAGGAGATC 1233
DB 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
QY 1234 ATAGATGTCATAGATAAGGAACATAATTCAACCTTTTGGAGATTCTATTGGAAGGAGTTGAA 1293
DB 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
QY 1294 TATATTGCCAGACTGGAGTCTCTGAGGGAATAATATGCTTGGTCCATCTACTAGAT 1353
DB 361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
QY 1354 CGCTCCAGACTCGCTACAGATAGTGTGATCTCCCTGAATTTATTTATCCAGTAGAA 1413
DB 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
QY 1414 CATGATGTTATGAAAGGCGAGAGACTCATTTGAGTCAGTCCCTGATCTTGTGACGCCACTA 1473
DB 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAepSerValThrProLeu 420
QY 1474 ATTATCTATGAAGAAACACAGACATCTCGATAAATATCCATGACATCTTTTCATGTTT 1533
DB 421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
QY 1534 CCCCAAAGTCACGAAGGAAATTTGAGTTTATTTTGGCTCTGAAATCAAAACAGGTTTC 1593
DB 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
QY 1594 CGTCATTTATACAAAATTACATCTATTTTAAAGGAAAGCAAAATATAAACCGATCCAGTGGT 1653
DB 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
QY 1654 GGGCTGCTGCTCCAGTGTATTTCAAGTGTCTCTATCAAGAGAGAGATAGCAATTTACCAAGT 1713
DB 481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluIleAlaIleThrSer 500
QY 1714 GGTCAATGGGAAGTCTTGGCCGCGCATGCTTAATATCCAGTTGATGAAGTCAGAAGG 1773
DB 501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520
QY 1774 CTGCTATATTTTGAAGCCCAAGACTCCCTTTAGAGCATCACTCTAGTGTAGTCAAGT 1833
DB 521 LeuValTyrPheGluGlyThrLysAepSerProLeuGluHisLeuTyrValIleSer 540
QY 1834 TACGTAAATCTCGAGAGGTGACAGCTGACTGACCGTGGCTACTCATCTTCTGCTGC 1893
DB 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgglyTyrSerHisSerCysCys 560
QY 1894 ATCAGTCAGACTGTGACTTTCTTTATAGTATAGTAACCAAGAAATCCACACTGT 1953
DB 561 IleSerGlnHisCysAepPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
QY 1954 GTGTCCTCTTACAGCTTATCAAGTCTCTGAAGATGACCCCAACTGTGCAAAACAAAGAAATT 2013
DB 581 ValSerLeuTyrLysLeuSerSerProGluAspAepProThrCysLysThrLysGluPhe 600
```



```
QY 2014 TGGCCACCAATTTGGATTGAGGCTCTCTCTGACTACTACTCTCTCCAGAAATTTTC 2073
Db 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe 620
QY 2074 TCTTTGAAAGTACTACTGATTTACATTTGATGGATGCTCTCAAGCCCTCATGACTTA 2133
Db 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
QY 2134 CAGCTCGAAGAAATATCTACTGTGCTGTTCATATATGTTGGTCTCTCAGGTGCACTTG 2193
Db 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu 660
QY 2194 GTGAATAATCGGTTTAAAGAGTCAAGTATTTCCGCTTTGAATACCCCTAGCTCTCTAGGT 2253
Db 661 ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly 680
QY 2254 TATGTGTTGTAGTATAGAACACAGGGGATCCTGTACCCAGGGGCTTAATTTGAGGCC 2313
Db 681 TyrValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly 700
QY 2314 GCCTTTAAATATAAATGGGTCAATAGAAATTCAGGATCAGGTGGAAGGACTCCAATAT 2373
Db 701 AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr 720
QY 2374 CTAGCTTCTCGATATGATTTCATTTGACTTATGATCGTGTGGGATCCAGGCTGCTCTAT 2433
Db 721 LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrSerTyr 740
QY 2434 GGAGGATACCTCTCCCTGATGGCAATTAATGCAGAGTGCAGATATCTTCAGGGTTCCTATT 2493
Db 741 GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760
QY 2494 GCTGGGCCCCAGTCACTCTGTGATCTTCTATGATACAGGATACAGGAACGTTATATG 2553
Db 761 AlaGlyAlaProValThrLeuThrIlePheTyrAspThrGlyTyrThrGluArgTyrMet 780
QY 2554 GGTCAACCTGACAGAAATGAAAGGCTATTACTTAGGATCTGTGGCCATCCAGGACAGAA 2613
Db 781 GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu 800
QY 2614 AAGTTCCTCTGAAACCAATCGTTTACTGCTCTTACATGTTTCCTGGATGAGAAATCTC 2673
Db 801 LysPheProSerGluProAsnArgLeuLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820
QY 2674 CATTTTCACATACCACTATATTACTAGTGTGTTTTAGTGAGGCTGGAAGCCATATGAT 2733
Db 821 HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840
QY 2734 TTACAGATCTATCTCTCAGGAGACACAGCATAAGAGTTTCTTGATCGGGAGAACATTAT 2793
Db 841 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860
QY 2794 GAATGATCTTTTGCACTACCTTCAAGAAACCTTGGATCAGTATTGCTGCTCAAAA 2853
Db 861 GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeuLys 880
QY 2854 GTGATA 2859
Db 881 ValIle 882

RESULT 2
US-11-176-951-10
; Sequence 10, Application US/11176951
; Publication No. US20060024313A1
; GENERAL INFORMATION:
; APPLICANT: CHEN, YUAN-SHOU
; TITLE OF INVENTION: AGENTS THAT DISRUPT DIMER FORMATION IN DPP-IV FAMILY OF
; FILE REFERENCE: 08842.0019
; CURRENT FILING DATE: 2005-07-06
; PRIOR APPLICATION NUMBER: 60/586,095
; PRIOR FILING DATE: 2004-07-06
; PRIOR APPLICATION NUMBER: 60/585,952
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 16
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-176-951-10
Alignment Scores:
Pred. No.: 8,02e-16 Length: 24
Score: 24.00 Matches: 24
```

```
; PRIOR APPLICATION NUMBER: 60/586,095
; PRIOR FILING DATE: 2004-07-06
; PRIOR APPLICATION NUMBER: 60/585,952
; PRIOR FILING DATE: 2004-07-06
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 10
; LENGTH: 109
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-176-951-10
Alignment Scores:
Pred. No.: 4.87e-104 Length: 109
Score: 109.00 Matches: 109
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 10.6% Indels: 0
Db: 11 Gaps: 0

US-10-825-632-2 (1-3120) x US-11-176-951-10 (1-109)
QY 2533 GGATACACGGAACGTTATATGGGTCACTGACCAAGATGAACAGGGCTATTACTTAGGA 2592
Db 1 GlyTyrThrGluArgTyrMetGlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGly 20
QY 2593 TCTGTGCCATCAAGCAGCAAAAGTTCCCTCTCTGAAACCAATCGTTTACTGCTCTTACAT 2652
Db 21 SerValAlaMetGlnAlaGluLysPheProSerGluProAsnArgLeuLeuLeuHis 40
QY 2653 GGTTCCTCGATGAGATGTCATTTTGCACATACCATATATTACTGATTTTACTGTG 2712
Db 41 GlyPheLeuAspGluAsnValHisPheAlaHisThrSerIleLeuLeuSerPheLeuVal 60
QY 2713 AGGCTCGAAAGCCATATGATTTTACAGATCTATCTCTCAGGAGACACAGCATAAGAGTT 2772
Db 61 ArgAlaGlyLysProTyrAspLeuGlnIleTyrProGlnGluArgHisSerIleArgVal 80
QY 2773 CCTGAATCGGAGAACATTAATGAATGCACTCTTTTGCATCTTCTCAAGAAACCTTGA 2832
Db 81 ProGluSerGlyGluHisTyrGluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGly 100
QY 2833 TCAGTATTGCTGCTCTTAAAGTGATA 2859
Db 101 SerArgIleAlaAlaLeuLysValIle 109

RESULT 3
US-11-176-951-16
; Sequence 16, Application US/11176951
; Publication No. US20060024313A1
; GENERAL INFORMATION:
; APPLICANT: CHEN, XIN
; APPLICANT: CHEN, YUAN-SHOU
; TITLE OF INVENTION: AGENTS THAT DISRUPT DIMER FORMATION IN DPP-IV FAMILY OF
; FILE REFERENCE: 08842.0019
; CURRENT FILING DATE: 2005-07-06
; PRIOR APPLICATION NUMBER: 60/586,095
; PRIOR FILING DATE: 2004-07-06
; PRIOR APPLICATION NUMBER: 60/585,952
; PRIOR FILING DATE: 2004-07-06
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 16
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-176-951-16
Alignment Scores:
Pred. No.: 8,02e-16 Length: 24
Score: 24.00 Matches: 24
```


NAME/KEY: SITE
LOCATION: (193)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-11-264-096-1593

Alignment Scores:
Pred. No.: 0.0241 Length: 497
Score: 11.00 Matches: 11
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 1.1% Indels: 0
DB: 11 Gaps: 0

US-10-825-632-2 (1-3120) x US-11-264-096-1593 (1-497)

QY 1804 CCTTGTAGCATCACCTGTACGTAGTCAGTTAC 1836
Db 156 ProluGluHisLeuTyValValSerTy 166

RESULT 7

US-11-264-096-1594
Sequence 1594, Application US/11264096
Publication No. US20060084794A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Albumin Fusion Proteins
FILE REFERENCE: PF546D1
CURRENT APPLICATION NUMBER: US/11/264,096
CURRENT FILING DATE: 2005-11-02
PRIOR APPLICATION NUMBER: 09/833,245
PRIOR FILING DATE: 2001-04-12
PRIOR APPLICATION NUMBER: 60/229,358
PRIOR FILING DATE: 2000-04-12
PRIOR APPLICATION NUMBER: 60/256,931
PRIOR FILING DATE: 2000-12-21
PRIOR APPLICATION NUMBER: 60/199,384
PRIOR FILING DATE: 2000-04-25
NUMBER OF SEQ ID NOS: 2267
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1594
LENGTH: 497
TYPE: PRT
ORGANISM: Homo sapiens
US-11-264-096-1594

Alignment Scores:
Pred. No.: 0.0241 Length: 497
Score: 11.00 Matches: 11
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 1.1% Indels: 0
DB: 11 Gaps: 0

US-10-825-632-2 (1-3120) x US-11-264-096-1594 (1-497)

QY 1804 CCTTGTAGCATCACCTGTACGTAGTCAGTTAC 1836
Db 156 ProluGluHisLeuTyValValSerTy 166

RESULT 8

US-11-172-740-719
Sequence 719, Application US/11172740
Publication No. US20060057724A1
GENERAL INFORMATION:
APPLICANT: MASCIA, Peter
APPLICANT: ALEXANDROV, Nickolai
APPLICANT: BROVER, Vyacheslav
TITLE OF INVENTION: NUCLEOTIDE SEQUENCES AND POLYPEPTIDES ENCODED THEREBY USEFUL FOR
TITLE OF INVENTION: PLANT CHARACTERISTICS AND PHENOTYPES
FILE REFERENCE: 2750-1602PUS2
CURRENT APPLICATION NUMBER: US/11/172,740
CURRENT FILING DATE: 2005-06-30
PRIOR APPLICATION NUMBER: 60/583,621

PRIOR FILING DATE: 2004-06-30
PRIOR APPLICATION NUMBER: 60/584,829
PRIOR FILING DATE: 2004-06-30
PRIOR APPLICATION NUMBER: 60/584,800
PRIOR FILING DATE: 2004-06-30
NUMBER OF SEQ ID NOS: 2523
SEQ ID NO 719
LENGTH: 353
TYPE: PRT
ORGANISM: Arabidopsis thaliana
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)-(353)
OTHER INFORMATION: Public GI no. 50906397
FEATURE:
NAME/KEY: misc_feature
LOCATION:
LOCATION:
OTHER INFORMATION: Utility: Useful for delaying flowering time
FEATURE:
NAME/KEY: misc_feature
LOCATION:
LOCATION:
OTHER INFORMATION: Utility: Useful for increasing chlorophyll and photosynthetic cap.
FEATURE:
NAME/KEY: misc_feature
LOCATION:
LOCATION:
OTHER INFORMATION: Utility: Useful for increasing anthocyanin content
FEATURE:
NAME/KEY: misc_feature
LOCATION:
LOCATION:
OTHER INFORMATION: Utility: Useful for making plants with increased biomass and foli.
NAME/KEY: misc_feature
LOCATION:
LOCATION:
OTHER INFORMATION: Utility: Useful for making smaller plants
US-11-172-740-719

Alignment Scores:
Pred. No.: 2.88 Length: 353
Score: 9.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 0.9% Indels: 0
DB: 11 Gaps: 0

US-10-825-632-2 (1-3120) x US-11-172-740-719 (1-353)

QY 23 AAGCGCTGCTACTGCGCGCTGCT 49
Db 313 LysalaalaThralaAlaAla 321

RESULT 9

US-11-071-259-3
Sequence 3, Application US/11071259
Publication No. US2006008440A1
GENERAL INFORMATION:
APPLICANT: Blatt, Thomas
APPLICANT: Mummert, Christopher
APPLICANT: Mundt, Claudia
APPLICANT: Staeb, Franz
TITLE OF INVENTION: Cosmetic or Dermatological Preparations Having a Content of
TITLE OF INVENTION: Anti-Freezing Proteins and/or Anti-Freezing Glycoproteins
FILE REFERENCE: P26688
CURRENT APPLICATION NUMBER: US/11/071,259
CURRENT FILING DATE: 2005-03-04
PRIOR APPLICATION NUMBER: PCT/EP2002/010044
PRIOR FILING DATE: 2002-09-07
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.3
SEQ ID NO 3
LENGTH: 49
TYPE: PRT
ORGANISM: Limanda ferruginea
US-11-071-259-3

Alignment Scores:
Pred. No.: 32 Length: 49
Score: 8.00 Matches: 8
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 0.8% Indels: 0
DB: 11 Gaps: 0

US-10-825-632-2 (1-3120) x US-11-071-259-3 (1-49)

Qy 26 GCCTGCTACTGCGCGCTGCT 49
Db 10 AlaAlaAlaThraAlaAlaAla 17

RESULT 10

US-11-096-568A-9319
; Sequence 9319, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:

; APPLICANT: Alexandrov, Nikolai et al.

; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides

; FILE REFERENCE: 2750-1592PUS2

; CURRENT APPLICATION NUMBER: US/11/096,568A

; CURRENT FILING DATE: 2005-04-01

; NUMBER OF SEQ ID NOS: 34471

; SEQ ID NO 9319

; LENGTH: 90

; TYPE: PRT

; ORGANISM: Triticum aestivum

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: (1)..(90)

; OTHER INFORMATION: Ceres Seq. ID no. 15226318

US-11-096-568A-9319

Alignment Scores:
Pred. No.: 31.8 Length: 90
Score: 8.00 Matches: 8
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 0.8% Indels: 0
DB: 11 Gaps: 0

US-10-825-632-2 (1-3120) x US-11-096-568A-9319 (1-90)

Qy 48 GCACGGCGGCGAGTACGACGCGC 25
Db 45 AlaAlaAlaValaAlaAlaAla 52

RESULT 11

US-10-504-588-6

; Sequence 6, Application US/10504588

; Publication No. US20060078498A1

; GENERAL INFORMATION:

; APPLICANT: Buckholz, Thomas

; APPLICANT: Vandenberg, Mark

; APPLICANT: Heitmeier, Stefan

; APPLICANT: Taylor, Ian

; APPLICANT: Gedrich, Richard

; TITLE OF INVENTION: Methods for the Identification of Novel Ligands for the G

; FILE REFERENCE: 5121-PCT

; CURRENT APPLICATION NUMBER: US/10/504,588

; CURRENT FILING DATE: 2004-08-12

; PRIOR FILING DATE: 2003-03-28

; PRIOR APPLICATION NUMBER: PCT/US03/09522

; PRIOR FILING DATE: 2003-03-28

; PRIOR APPLICATION NUMBER: US 60/368,849

; NUMBER OF SEQ ID NOS: 18

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 6

; LENGTH: 108
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-504-588-6
Alignment Scores:
Pred. No.: 31.8 Length: 108
Score: 8.00 Matches: 8
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 0.8% Indels: 0
DB: 9 Gaps: 0

US-10-825-632-2 (1-3120) x US-10-504-588-6 (1-108)

Qy 30 CTGCTACTGCGCGCTGCTTCTT 53
Db 13 LeuLeuLeuProLeuLeuLeu 20

RESULT 12

US-11-073-420-6

; Sequence 6, Application US/11073420

; Publication No. US20060019338A1

; GENERAL INFORMATION:

; APPLICANT: Zhou, Qun-Yong

; TITLE OF INVENTION: Primate Prokineticin and Prokineticin

; FILE REFERENCE: UC11210-1

; CURRENT APPLICATION NUMBER: US/11/073,420

; CURRENT FILING DATE: 2005-03-04

; PRIOR APPLICATION NUMBER: 60/550,753

; PRIOR FILING DATE: 2004-03-05

; NUMBER OF SEQ ID NOS: 38

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 6

; LENGTH: 108

; TYPE: PRT

; ORGANISM: Macaca mulatta

; FEATURE:

; NAME/KEY: SIGNAL

; LOCATION: (1)..(27)

US-11-073-420-6

Alignment Scores:
Pred. No.: 31.8 Length: 108
Score: 8.00 Matches: 8
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 0.8% Indels: 0
DB: 11 Gaps: 0

US-10-825-632-2 (1-3120) x US-11-073-420-6 (1-108)

Qy 30 CTGCTACTGCGCGCTGCTTCTT 53
Db 13 LeuLeuLeuProLeuLeuLeu 20

RESULT 13

US-11-194-246-399

; Sequence 399, Application US/11194246

; Publication No. US20050272089A1

; GENERAL INFORMATION:

; APPLICANT: Mott, John

; APPLICANT: Trepod, Catherine

; APPLICANT: Arvidson, Staifan

; TITLE OF INVENTION: CRITICAL GENES AND POLYPEPTIDES OF HAEMOPHILUS INFLUENZAE AND METT

; FILE REFERENCE: 00592.US1 (M&R 268.05920101)

; CURRENT APPLICATION NUMBER: US/11/194,246

; CURRENT FILING DATE: 2005-08-01

; PRIOR APPLICATION NUMBER: US/10/274,586

; PRIOR FILING DATE: 2002-10-21

;; PRIOR APPLICATION NUMBER: US 60/345,438
;; PRIOR FILING DATE: 2001-10-19
;; NUMBER OF SEQ ID NOS: 621
;; SOFTWARE: PatentIn version 3.0
;; SEQ ID NO 399
;; LENGTH: 123
;; TYPE: PRT
;; ORGANISM: HAEMOPHILUS INFLUENZAE
US-11-194-246-399

Alignment Scores:
Pred. No.: 31.7 Length: 123
Score: 8.00 Matches: 8
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 0.8% Indels: 0
DB: 11 Gaps: 0

US-10-825-632-2 (1-3120) x US-11-194-246-399 (1-123)

Qy 48 GCAGCGCGGCAGTAGCAGCGGCC 25
Db 35 AlaAlaAlaValAlaAlaAla 42

RESULT 14

US-11-096-568A-14806
; Sequence 14806, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Therby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 14806
; LENGTH: 129
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(129)
; OTHER INFORMATION: Ceres Seq. ID no. 11415220
US-11-096-568A-14806

Alignment Scores:
Pred. No.: 31.7 Length: 129
Score: 8.00 Matches: 8
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 0.8% Indels: 0
DB: 11 Gaps: 0

US-10-825-632-2 (1-3120) x US-11-096-568A-14806 (1-129)

Qy 48 GCAGCGCGGCAGTAGCAGCGGCC 25
Db 5 AlaAlaAlaValAlaAlaAla 12

RESULT 15

US-11-096-568A-26015
; Sequence 26015, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Therby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 26015

;; LENGTH: 129
;; TYPE: PRT
;; ORGANISM: Zea mays subsp. mays
;; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(129)
; OTHER INFORMATION: Ceres Seq. ID no. 13498137
US-11-096-568A-26015

Alignment Scores:
Pred. No.: 31.7 Length: 129
Score: 8.00 Matches: 8
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 0.8% Indels: 0
DB: 11 Gaps: 0

US-10-825-632-2 (1-3120) x US-11-096-568A-26015 (1-129)

Qy 48 GCAGCGCGGCAGTAGCAGCGGCC 25
Db 5 AlaAlaAlaValAlaAlaAla 12

Search completed: May 2, 2006, 05:01:48
Job time : 69.5 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 4, 2006, 03:41:02 ; Search time 1064.84 Seconds
(without alignments)

Title: US-10-825-632-4

Perfect score:

Sequence: 1 atttgaaggcaccaagac.....ttactaaaaaaaaaaaaa 1197

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

```

Database : Published_Applications_NA_Main:*
1: /cgn2_6/pdata/1/pubnpa/US07_PUBCOMB.seq:*
2: /cgn2_6/pdata/1/pubnpa/US08_PUBCOMB.seq:*
3: /cgn2_6/pdata/1/pubnpa/US09A_PUBCOMB.seq:*
4: /cgn2_6/pdata/1/pubnpa/US09B_PUBCOMB.seq:*
5: /cgn2_6/pdata/1/pubnpa/US10A_PUBCOMB.seq:*
6: /cgn2_6/pdata/1/pubnpa/US10B_PUBCOMB.seq:*
7: /cgn2_6/pdata/1/pubnpa/US10C_PUBCOMB.seq:*
8: /cgn2_6/pdata/1/pubnpa/US10D_PUBCOMB.seq:*
9: /cgn2_6/pdata/1/pubnpa/US10E_PUBCOMB.seq:*
10: /cgn2_6/pdata/1/pubnpa/US11_PUBCOMB.seq:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	1197	100.0	1197	8	US-10-825-632-4	Sequence 4, Appli
2	1179	98.5	4676	3	US-09-976-674-20	Sequence 20, Appl
3	1179	98.5	4676	3	US-10-982-513-20	Sequence 20, Appl
4	1040	86.9	3143	6	US-10-170-789-37	Sequence 37, Appl
5	1036	86.5	3120	7	US-10-415-123-5	Sequence 5, Appl
6	1036	86.5	3120	8	US-10-825-632-2	Sequence 2, Appli
7	1033.2	86.3	4309	3	US-09-976-674-14	Sequence 14, Appl
8	1033.2	86.3	4309	9	US-10-982-513-14	Sequence 14, Appl
9	1033.2	86.3	4829	3	US-09-976-674-12	Sequence 12, Appl
10	1033.2	86.3	4829	3	US-10-982-513-12	Sequence 12, Appl
11	1033	86.3	3106	7	US-10-311-035-30	Sequence 30, Appl
12	1013.6	84.7	4685	3	US-09-976-674-22	Sequence 22, Appl
13	1013.6	84.7	4685	9	US-10-982-513-22	Sequence 22, Appl
14	881	73.6	1669	8	US-10-825-633-6	Sequence 6, Appli
15	881	73.6	2830	9	US-10-956-157-2177	Sequence 2177, Ap
16	873	72.9	4523	3	US-09-976-674-8	Sequence 8, Appli
17	873	72.9	4523	9	US-10-982-513-8	Sequence 8, Appli
18	863.6	72.1	2510	7	US-10-975-505-16	Sequence 16, Appl
19	863.6	72.1	2510	10	US-11-140-224-16	Sequence 16, App
20	791.4	66.1	925	6	US-10-264-237-710	Sequence 710, App
21	778	65.0	2649	6	US-10-170-789-39	Sequence 39, Appl
22	776.4	64.9	2649	6	US-10-054-776-1	Sequence 1, Appli
23	776.4	64.9	2671	3	US-09-976-674-2	Sequence 2, Appli

Sequence 2, Appli
Sequence 7412, Appli
Sequence 8, Appli
Sequence 877 Appli
Sequence 19585, A
Sequence 3, Appli
Sequence 223, Appli
Sequence 225, Appli
Sequence 8, Appli
Sequence 4, Appli
Sequence 4, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 30, Appli
Sequence 36, Appli
Sequence 36, Appli
Sequence 28, Appli
Sequence 28, Appli
Sequence 34, Appli
Sequence 34, Appli
Sequence 24, Appli
Sequence 24, Appli
Sequence 47, Appli

ALIGNMENTS

```

RESULT 1
US-10-825-632-4
; Sequence 4, Application US/10825632
; Publication No. US20040191826A1
; GENERAL INFORMATION:
; APPLICANT: ABBOTT, Catherine Anne
; APPLICANT: GORRELL, Mark Douglas
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES
; FILE REFERENCE: FCSB-100-Div. 1
; CURRENT APPLICATION NUMBER: US/10/825,632
; CURRENT FILING DATE: 2004-04-15
; PRIOR APPLICATION NUMBER: US 10/070,464
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: PCT/AU00/01085
; PRIOR FILING DATE: 2000-09-11
; PRIOR APPLICATION NUMBER: AU PQ5709
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: AU PQ2762
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1197

```

Query Match	100.0%	Score 1197;	DB 8;	Length 1197;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1197: Conservative	0;	Mismatches	0;	Indels 0;
Gaps	0;			

Qy	1	ATTTTGAAGCCACCAAGACTCCCTTTT	AGAGCATCACCTGTACGTAGTCAGTTAGCTAA	60
Db	1	ATTTTGAAGGCACCAAGACTCCCTTT	AGAGCATCACCTGTACGTAGTCAGTTAGCTAA	60
Qy	61	ATCTCGAGAGGTGACAAGGCTGACTGAC	CGTGCTACTCACATTTCTTGTGCATCAGTC	120
Db	61	ATCTCGAGAGGTGACAAGGCTGACTGAC	CGTGCTACTCACATTTCTTGTGCATCAGTC	120
Qy	121	AGCACTGTGACTTCTTTTATAAGTAAGTA	TAGTAACCAAGAGAATCCACACTGTGTGCTCC	180
Db	121	AGCACTGTGACTTCTTTTATAAGTAAGTA	TAGTAACCAAGAGAATCCACACTGTGTGCTCC	180
Qy	181	TTTACAAGCTATCAAGTCCTCAAGATGACC	CCCACTTGCAAAACAAAGGAATTTTGGGCCA	240
Db	181	TTTACAAGCTATCAAGTCCTCAAGATGACC	CCCACTTGCAAAACAAAGGAATTTTGGGCCA	240

QY 241 CCATTTGGATTACAGAGGCTCCCTTCTGACTATATCTCTCCAGAAATTTCTCTTTG 300
Db 241 CCATTTGGATTACAGAGGCTCCCTTCTGACTATATCTCTCCAGAAATTTCTCTTTG 300
QY 301 AAAGTACTACTGGATTTACATTTGATGGATGCTCTACAGGCTCATGATCTACAGGCTG 360
Db 301 AAAGTACTACTGGATTTACATTTGATGGATGCTCTACAGGCTCATGATCTACAGGCTG 360
QY 361 GAAGAATAATCTCTACTGTGCTGTTTCATATATGATGGTGGTCTCAGGGTCAAAATAGAAATG 420
Db 361 GAAGAATAATCTCTACTGTGCTGTTTCATATATGATGGTGGTCTCAGGGTCAAAATAGAAATG 420
QY 421 ACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTTCACTTAGATC 480
Db 421 ACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTTCACTTAGATC 480
QY 481 GTGTGGGCATCCACGGCTGTCTCTATGGAGGATACCTCTCCCTGATGGCAATTAATGAGA 540
Db 481 GTGTGGGCATCCACGGCTGTCTCTATGGAGGATACCTCTCCCTGATGGCAATTAATGAGA 540
QY 541 GGTGAGATATCTTCAGGGTTGCTATTGCTGGGGGCCAGTCACTCTGTGGATCTTCTATG 600
Db 541 GGTGAGATATCTTCAGGGTTGCTATTGCTGGGGGCCAGTCACTCTGTGGATCTTCTATG 600
QY 601 ATACAGGATACACGGAACGTTATATGGTCAACCTGACAGAAATGAACAGGCTATTACT 660
Db 601 ATACAGGATACACGGAACGTTATATGGTCAACCTGACAGAAATGAACAGGCTATTACT 660
QY 661 TAGGATCTGTGGCCATCGAAGCAGAAAGTTCCCTCTGAACCAATCGTTTACTGCTCT 720
Db 661 TAGGATCTGTGGCCATCGAAGCAGAAAGTTCCCTCTGAACCAATCGTTTACTGCTCT 720
QY 721 TACATGGTTTCTCGATGAGAAATGTCATTTTGACATACCAAGTATATTACTAGTTT 780
Db 721 TACATGGTTTCTCGATGAGAAATGTCATTTTGACATACCAAGTATATTACTAGTTT 780
QY 781 TAGTGAGGCTGGAAGCCATATGATTTACAGATCTATCTCTCAGAGAGACACAGCATAA 840
Db 781 TAGTGAGGCTGGAAGCCATATGATTTACAGATCTATCTCTCAGAGAGACACAGCATAA 840
QY 841 GAGTTCCTGAATCGGGAGAACATTTATGAATGATCTTTTGCACTACCTTCAAGAAACC 900
Db 841 GAGTTCCTGAATCGGGAGAACATTTATGAATGATCTTTTGCACTACCTTCAAGAAACC 900
QY 901 TTGGATCACGTTATTTGCTGCTTAAAGTGATATAATTTTGACCTGTGTAGAACTCTCTG 960
Db 901 TTGGATCACGTTATTTGCTGCTTAAAGTGATATAATTTTGACCTGTGTAGAACTCTCTG 960
QY 961 TATACACTGGCTATTTAAACCAATGAGGAGTTTAAATCAACAGAAACACAGAAATGATC 1020
Db 961 TATACACTGGCTATTTAAACCAATGAGGAGTTTAAATCAACAGAAACACAGAAATGATC 1020
QY 1021 ATCATTATTTGATACCTGCCATGTAAATCTACTCTGAAATAAATGTTGGTGCATGCA 1080
Db 1021 ATCATTATTTGATACCTGCCATGTAAATCTACTCTGAAATAAATGTTGGTGCATGCA 1080
QY 1081 GGGTCTACGGTTTGTGGTAGTAATCTAATACCTTAACCCACATGCTCAAAATCAAAATG 1140
Db 1081 GGGTCTACGGTTTGTGGTAGTAATCTAATACCTTAACCCACATGCTCAAAATCAAAATG 1140
QY 1141 ATACATATCTCTGAGAGACCCAGCAATACCAATAGAAATTTACTAAAAA 1197
Db 1141 ATACATATCTCTGAGAGACCCAGCAATACCAATAGAAATTTACTAAAAA 1197

RESULT 2
US-09-976-674-20
; Sequence 20, Application US/09976674
; Patent No. US20020115843A1
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen

; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 20
; LENGTH: 4676
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-976-674-20

Query Match 98.5%; Score 1179; DB 3; Length 4676;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1193; Conservative 0; Mismatches 0; Indels 4; Gaps 1;

QY 1 ATTTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTACGTAGCTTACGTAA 60
Db 1781 ATTTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTACGTAGCTTACGTAA 1840
QY 61 ATCTGTGAGAGGTGACAAAGGCTGACTGACCGTGGCTACTCATTTCTTGTGCACTGTC 120
Db 1841 ATCTGTGAGAGGTGACAAAGGCTGACTGACCGTGGCTACTCATTTCTTGTGCACTGTC 1900
QY 121 AGCACTGTGACTTCTTTTATAGTAAGTATAGTAACCAAGAAATCCACACTGTGTGTCCTC 180
Db 1901 AGCACTGTGACTTCTTTTATAGTAAGTATAGTAACCAAGAAATCCACACTGTGTGTCCTC 1960
QY 181 TTTTCAAGCTATCAAGTCTGGAAGTACCAAGTGTGAAACCAAGAAATTTTGGGCA 240
Db 1961 TTTTCAAGCTATCAAGTCTGGAAGTACCAAGTGTGAAACCAAGAAATTTTGGGCA 2020
QY 241 CCATTTTGGATTCAGCAGGTCCTCTCTGACTATATCTCTCCAGAAATTTCTCTTTG 300
Db 2021 CCATTTTGGATTCAGCAGGTCCTCTCTGACTATATCTCTCCAGAAATTTCTCTTTG 2080
QY 301 AAAGTACTACTGGATTTACATTTGATGGATGCTCTACAGGCTCATGATCTACAGGCTG 360
Db 2081 AAAGTACTACTGGATTTACATTTGATGGATGCTCTACAGGCTCATGATCTACAGGCTG 2140
QY 361 GAAAGAAATATCTCTACTGTGCTGTTTCATATATGTTGGTGGTCTCAGGGTCAAAATAGAAATG 420
Db 2141 GAAAGAAATATCTCTACTGTGCTGTTTCATATATGTTGGTGGTCTCAGGGTCAAAATAGAAATG 2196
QY 421 ACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTTCACTTAGATC 480
Db 2197 ACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTTCACTTAGATC 2256
QY 481 GTGTGGGCATCCACGGCTGGTCTCTATGGAGGATACCTCTCCCTGATGGCAATTAATGAGA 540
Db 2257 GTGTGGGCATCCACGGCTGGTCTCTATGGAGGATACCTCTCCCTGATGGCAATTAATGAGA 2316
QY 541 GGTGAGATATCTTCAGGGTTGCTATTGCTGGGGGCCAGTCACTCTGTGGATCTTCTATG 600
Db 2317 GGTGAGATATCTTCAGGGTTGCTATTGCTGGGGGCCAGTCACTCTGTGGATCTTCTATG 2376
QY 601 ATACAGGATACCGGACGTTATATGGTCAACCTGACAGAAATGAACAGGCTATTACT 660
Db 2377 ATACAGGATACCGGACGTTATATGGTCAACCTGACAGAAATGAACAGGCTATTACT 2436
QY 661 TAGGATCTGTGGCCATGCAAGCAGAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCT 720
Db 2437 TAGGATCTGTGGCCATGCAAGCAGAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCT 2496
QY 721 TACATGGTTTCTCGATGAGAAATGTCATTTTGCAATACCAAGTATATTACTAGTTTTT 780
Db 2497 TACATGGTTTCTCGATGAGAAATGTCATTTTGCAATACCAAGTATATTACTAGTTTTT 2556
QY 781 TAGTGAGGCTGGAAGCCATATGATTTTACAGATCTATCTCTCAGGAGAGACACAGCATAA 840

Db 2557 TAGTGGGCTGGAAGCCATATGTTTACAGATCTATCTCAGGAGACACAGCATAA 2616
Qy 841 GAGTTCCTGAATCGGAGAACATATGAACTGTCATCTTTTGGCTACTCTTCAAGAAACC 900
Db 2617 GAGTTCCTGAATCGGAGAACATATGAACTGTCATCTTTTGGCTACTCTTCAAGAAACC 2676
Qy 901 TTGGATCAGGATTCGCTGCTCTAAAGTGATATAATTTTGCACCTGTGTAGAACTCTCTGG 960
Db 2677 TTGGATCAGGATTCGCTGCTCTAAAGTGATATAATTTTGCACCTGTGTAGAACTCTCTGG 2736
Qy 961 TATACACTGGCTATTTAAACCAATGAGGAGGTTTAAATCAACAGAAACACAGAAATGATC 1020
Db 2737 TATACACTGGCTATTTAAACCAATGAGGAGGTTTAAATCAACAGAAACACAGAAATGATC 2796
Qy 1021 ATCACAATTTTGATACCTGCGATGTAACATCTACTCTCGAAATAAATGTTGGTCCATGCA 1080
Db 2797 ATCACAATTTTGATACCTGCGATGTAACATCTACTCTCGAAATAAATGTTGGTCCATGCA 2856
Qy 1081 GGGGCTCTACGGTTTGTGGTAGTAATCTAATACCTTAAACCCACATGCTCAAAATCAAATG 1140
Db 2857 GGGGCTCTACGGTTTGTGGTAGTAATCTAATACCTTAAACCCACATGCTCAAAATCAAATG 2916
Qy 1141 ATACATATTCCTGAGAGACCCAGCAATACCAATAAGAAATTTACTAAAAAATAAAAAA 1197
Db 2917 ATACATATTCCTGAGAGACCCAGCAATACCAATAAGAAATTTACTAAAAAATAAAAAA 2973

RESULT 3

US-10-982-512-20
; Sequence 20, Application US/10982512
; Publication No. US20050059081A1
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Junien, Jean-Louis
; APPLICANT: Riviere, Pierre
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/10/982,512
; CURRENT FILING DATE: 2004-11-05
; PRIOR APPLICATION NUMBER: US/09/976,674
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20
; LENGTH: 4676
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-982-512-20

Query Match 98.5%; Score 1179; DB 9; Length 4676;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1193; Conservative 0; Mismatches 0; Indels 4; Gaps 1;
Qy 1 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACCTGTACGTAGTTACGTAA 60
Db 1781 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACCTGTACGTAGTTACGTAA 1840
Qy 61 ATCTGAGAGGTGACAGGCTGACTGACCGTGGTCTTCACTTTCTTGTGTCATCAGTC 120
Db 1841 ATCTGAGAGGTGACAGGCTGACTGACCGTGGTCTTCACTTTCTTGTGTCATCAGTC 1900
Qy 121 AGCACTGTGACTCTTTTATAGTAAGTATAGTAACCAAGAAATCCACACTGTGTGCC 180
Db 1901 AGCACTGTGACTCTTTTATAGTAAGTATAGTAACCAAGAAATCCACACTGTGTGCC 1960
Qy 181 TTTTCAAGCTATCAAGTCTCTGAAGATGACCCAAATTTGCAAAACAAAGGAATTTTGGGCA 240
Db 1961 TTTTCAAGCTATCAAGTCTCTGAAGATGACCCAAATTTGCAAAACAAAGGAATTTTGGGCA 2020

RESULT 4

US-10-170-789-37
; Sequence 37, Application US/10170789
; Publication No. US20030180930A1
; GENERAL INFORMATION:
; APPLICANT: Rachel E. Meyers
; APPLICANT: Olandt, Peter J.
; APPLICANT: Kapellier-Libermann, Rosana

Qy 241 CCATTTTGGATTACAGAGTCTCTCTCTGACTATATCTCTCCAGAAAATTTTCTCTTTTG 300
Db 2021 CCATTTTGGATTACAGAGTCTCTCTCTGACTATATCTCTCCAGAAAATTTTCTCTTTTG 2080
Qy 301 AAAGTACTACTGGATTACATTTGATGGATGCTCTACAAGCCTCATGATCTACAGCCTG 360
Db 2081 AAAGTACTACTGGATTACATTTGATGGATGCTCTACAAGCCTCATGATCTACAGCCTG 2140
Qy 361 GAAAGAAAATATCTACTCTGCTGTTTCATATATGTTGGTCTCTCAGGGTCAAAATGAAATTTG 420
Db 2141 GAAAGAAAATATCTACTCTGCTGTTTCATATATGTTGGTCTCTCAGGGTCAAAATGAAATTTG 2196
Qy 421 ACATCAGGTGGAAGGATCTCAATATCTAGCTTCTCGATATGATTTCAATGACTTATAGATC 480
Db 2197 ACATCAGGTGGAAGGATCTCAATATCTAGCTTCTCGATATGATTTCAATGACTTATAGATC 2256
Qy 481 GTGTGGGATCCACAGGCTGCTCTATGAGGATACCTCTCTCCCTGATGGCATTAATGAGCA 540
Db 2257 GTGTGGGATCCACAGGCTGCTCTATGAGGATACCTCTCTCCCTGATGGCATTAATGAGCA 2316
Qy 541 GGTTCAGATATCTTTCAGGGTTCATTTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATG 600
Db 2317 GGTTCAGATATCTTTCAGGGTTCATTTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATG 2376
Qy 601 ATACAGGATACAGGAACTGTTATGAGGTCACCTTGACAGCAATGAACAGGCTTATTAAT 660
Db 2377 ATACAGGATACAGGAACTGTTATGAGGTCACCTTGACAGCAATGAACAGGCTTATTAAT 2436
Qy 661 TAGGATCTGTGGCATGCAAGCAGAAAAGTTCCCTCTGAAACCAAAATGTTTACTGCTCT 720
Db 2437 TAGGATCTGTGGCATGCAAGCAGAAAAGTTCCCTCTGAAACCAAAATGTTTACTGCTCT 2496
Qy 721 TACATGGTTTCTCTGGATGAGAAATGTCCTTTTGGACATACCAAGTATATTAATGAGTTTTT 780
Db 2497 TACATGGTTTCTCTGGATGAGAAATGTCCTTTTGGACATACCAAGTATATTAATGAGTTTTT 2556
Qy 781 TAGTGGGGCTGGAAGGCAATGATTTACAGATCTATCTCTCAGGAGAGACACAGCATAA 840
Db 2557 TAGTGGGGCTGGAAGGCAATGATTTACAGATCTATCTCTCAGGAGAGACACAGCATAA 2616
Qy 841 GAGTTCCTGAACTCGGAGAACATTTAGTCACTCTTTTGGACATACCTTCAAGAAAACC 900
Db 2617 GAGTTCCTGAACTCGGAGAACATTTAGTCACTCTTTTGGACATACCTTCAAGAAAACC 2676
Qy 901 TTGGATCACGTAATTTGCTGCTCTAAAGTATATAATTTTGAACCTGTGTAGAACTCTCTGG 960
Db 2677 TTGGATCACGTAATTTGCTGCTCTAAAGTATATAATTTTGAACCTGTGTAGAACTCTCTGG 2736
Qy 961 TATACACTGGCTATTTAAACCAATGAGGAGTTTAAATCAACAGAAAACACAGAAATTTGATC 1020
Db 2737 TATACACTGGCTATTTAAACCAATGAGGAGTTTAAATCAACAGAAAACACAGAAATTTGATC 2796
Qy 1021 ATCACATTTTGTATACCTGCGATCTAACAATCTACTCTCTGAAAATAAATGTTGGTCCATGCA 1080
Db 2797 ATCACATTTTGTATACCTGCGATCTAACAATCTACTCTCTGAAAATAAATGTTGGTCCATGCA 2856
Qy 1081 GGGGCTCTACGGTTTGTGGTAGTAATCTAATACCTTAAACCCACATGCTCAAAATCAAATG 1140
Db 2857 GGGGCTCTACGGTTTGTGGTAGTAATCTAATACCTTAAACCCACATGCTCAAAATCAAATG 2916
Qy 1141 ATACATATTCCTGAGAGACCCAGCAATACATAAGAAATTTACTAAAAAATAAAAAA 1197
Db 2917 ATACATATTCCTGAGAGACCCAGCAATACATAAGAAATTTACTAAAAAATAAAAAA 2973

APPLICANT: Curtis , Rory A. J.
APPLICANT: Williamson, Mark
APPLICANT: Welch, Nadine
TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASE, PHOSPHATASE,
FILE OF INVENTION: AND PROTEASE FAMILY MEMBERS AND USES THEREOF
FILE REFERENCE: 10448-191001
CURRENT APPLICATION NUMBER: US/10/170,789
CURRENT FILING DATE: 2002-06-13
PRIOR APPLICATION NUMBER: US 09/797,039
PRIOR FILING DATE: 2001-02-28
PRIOR APPLICATION NUMBER: PCT/US01/06525
PRIOR FILING DATE: 2001-02-28
PRIOR APPLICATION NUMBER: US 60/186,061
PRIOR FILING DATE: 2000-02-29
PRIOR APPLICATION NUMBER: US 09/882,166
PRIOR FILING DATE: 2001-06-15
PRIOR APPLICATION NUMBER: PCT/US01/19269
PRIOR FILING DATE: 2001-06-15
PRIOR APPLICATION NUMBER: US 60/212,078
PRIOR FILING DATE: 2000-06-15
PRIOR APPLICATION NUMBER: US 09/934,406
PRIOR FILING DATE: 2001-08-21
PRIOR APPLICATION NUMBER: PCT/US01/26052
PRIOR FILING DATE: 2001-08-21
PRIOR APPLICATION NUMBER: US 60/226,740
PRIOR FILING DATE: 2000-08-21
PRIOR APPLICATION NUMBER: US 09/861,801
PRIOR FILING DATE: 2001-05-21
PRIOR APPLICATION NUMBER: PCT/US01/16549
PRIOR FILING DATE: 2001-05-21
PRIOR APPLICATION NUMBER: US 60/205,508
PRIOR FILING DATE: 2000-05-19
PRIOR APPLICATION NUMBER: US 09/801,267
PRIOR FILING DATE: 2001-03-06
PRIOR APPLICATION NUMBER: PCT/US01/07138
PRIOR FILING DATE: 2001-03-05
PRIOR APPLICATION NUMBER: US 60/187,454
PRIOR FILING DATE: 2000-03-07
PRIOR APPLICATION NUMBER: US 09/829,671
PRIOR FILING DATE: 2001-04-10
PRIOR APPLICATION NUMBER: PCT/US01/40483
PRIOR FILING DATE: 2000-04-11
PRIOR APPLICATION NUMBER: US 60/197,508
PRIOR FILING DATE: 2000-04-18
PRIOR APPLICATION NUMBER: US 09/961,721
PRIOR FILING DATE: 2001-09-24
PRIOR APPLICATION NUMBER: PCT/US01/29904
PRIOR FILING DATE: 2001-09-24
PRIOR APPLICATION NUMBER: US 60/235,023
PRIOR FILING DATE: 2000-09-25
PRIOR APPLICATION NUMBER: US 10/045,367
PRIOR FILING DATE: 2001-11-07
PRIOR APPLICATION NUMBER: US 60/246,561
PRIOR FILING DATE: 2000-11-07
PRIOR APPLICATION NUMBER: US 09/801,275
PRIOR FILING DATE: 2001-03-06
PRIOR APPLICATION NUMBER: PCT/US01/07074
PRIOR FILING DATE: 2001-03-05
PRIOR APPLICATION NUMBER: US 60/187,420
PRIOR FILING DATE: 2000-03-07
NUMBER OF SEQ ID NOS: 63
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 37
LENGTH: 3143
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (229)...(2874)
US-10-170-789-37

Query Match 86.9%; Score 1040; DB 6; Length 3143;
Best Local Similarity 89.1%; Pred. No. 8.6e-303;

	Matches 1197;	Conservative 0;	Mismatches 0;	Indels 147;	Gaps 1;
Qy 1	ATTTTGAAGGACCAAAAGACTCCCTTTTAGAGATCACCTGACGTAGCTAGCTTACGTAA	60			
Db 1796	ATTTTGAAGGACCAAAAGACTCCCTTTTAGAGATCACCTGACGTAGCTAGCTTACGTAA	1855			
Qy 61	ATCCTGGAGAGGTGACAAAGGCTGACCGTGGCTACTCACAATCTTGTGTCATCAGTC	120			
Db 1856	ATCCTGGAGAGGTGACAAAGGCTGACCGTGGCTACTCACAATCTTGTGTCATCAGTC	1915			
Qy 121	AGCACTGTGACTTCTTTTAAAGTAAGTATAGTAAACAGAAAGATCCACACTGTGTGTC	180			
Db 1916	AGCACTGTGACTTCTTTTAAAGTAAGTATAGTAAACAGAAAGATCCACACTGTGTGTC	1975			
Qy 181	TTTAAAGCTTATCAAGTCTGAAAGATGACCAACTTGAACAAAGAAATTTTGGGCCA	240			
Db 1976	TTTAAAGCTTATCAAGTCTGAAAGATGACCAACTTGAACAAAGAAATTTTGGGCCA	2035			
Qy 241	CCATTTTGGATTCCAGCAGGTCTCTTCTGACTATCTCTCAAGCCTCATGCTCAGCCTG	300			
Db 2036	CCATTTTGGATTCCAGCAGGTCTCTTCTGACTATCTCTCAAGCCTCATGCTCAGCCTG	2095			
Qy 301	AAAGTACTACTGGATTACATTTGATGGATGCTCTCAAGCCTCATGCTCAGCCTG	360			
Db 2096	AAAGTACTACTGGATTACATTTGATGGATGCTCTCAAGCCTCATGCTCAGCCTG	2155			
Qy 361	GAAGAAATATCTCTACTGTGCTGTTTCATATATGTTGGTCTCTCA	403			
Db 2156	GAAGAAATATCTCTACTGTGCTGTTTCATATATGTTGGTCTCTCA	2215			
Qy 404	-----	403			
Db 2216	ATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCTAGCCTCTCTAGGTTATGTG	2275			
Qy 404	-----	403			
Db 2276	TTGTAGTGATAGACAAACAGGGGATCCTGTCCACGAGGCTTTAAATTTGAAGGCGCCTTA	2335			
Qy 404	-----GGGTCAAATAGAAAATTGACGATCAGGTGGAAGGACTCCAATATCTAGCTT	453			
Db 2336	AAATATAAAATGGGTCAAATAGAAAATTGACGATCAGGTGGAAGGACTCCAATATCTAGCTT	2395			
Qy 454	CTCGATATGATTTTCAATTTGACTTAGATCGTGTGGCATCCACGGTGTCTCTATGGAGGAT	513			
Db 2396	CTCGATATGATTTTCAATTTGACTTAGATCGTGTGGCATCCACGGTGTCTCTATGGAGGAT	2455			
Qy 514	ACCTCTCCCTGATGGCATTAAATGCAGAGTTCAGATATCTTCAGGGTTGCTATTGCTGGGG	573			
Db 2456	ACCTCTCCCTGATGGCATTAAATGCAGAGTTCAGATATCTTCAGGGTTGCTATTGCTGGGG	2515			
Qy 574	CCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGACGTTATATGGGTCAAC	633			
Db 2516	CCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGACGTTATATGGGTCAAC	2575			
Qy 634	CTGACAGAAATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAAAAGTTCC	693			
Db 2576	CTGACAGAAATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAAAAGTTCC	2635			
Qy 694	CCTCTGAACCAAAATCGTTTACTGTCTTACATGATGTTTCTGGATGAGAATGTCATTTTG	753			
Db 2636	CCTCTGAACCAAAATCGTTTACTGTCTTACATGATGTTTCTGGATGAGAATGTCATTTTG	2695			
Qy 754	CACATACCAGTATATTAAGTATTTTGTAGTGGGCTGGAAGCCATATGATTATACAGA	813			
Db 2696	CACATACCAGTATATTAAGTATTTTGTAGTGGGCTGGAAGCCATATGATTATACAGA	2755			
Qy 814	TCTATCTCTCAGGAGACACAGCAATAGAGTTCTGAATCGGGAGAACATTATGAATC	873			
Db 2756	TCTATCTCTCAGGAGACACAGCAATAGAGTTCTGAATCGGGAGAACATTATGAATC	2815			
Qy 874	ATCTTTTGTCACTACCTTCAAGAAAACCTTGGATCAGCTATTGCTCTCTAAAGTGATAT	933			
Db 2816	ATCTTTTGTCACTACCTTCAAGAAAACCTTGGATCAGCTATTGCTCTCTCTAAAGTGATAT	2875			

QY 934 AATTTTACCTGTGTAGAACTCTCTGGTATACACTGGCTATTAAACCAATGAGGAGTT 993
DB 2876 AATTTTACCTGTGTAGAACTCTCTGGTATACACTGGCTATTAAACCAATGAGGAGTT 2935
QY 994 TAATCAACAGAAACACAGAAATGATCATCATATTTTGTATACCTGCCATGTAAACATCTAC 1053
DB 2936 TAATCAACAGAAACACAGAAATGATCATCATATTTTGTATACCTGCCATGTAAACATCTAC 2995
QY 1054 TCCTGAAATAAATGTGGTGCCATCGAGGGGTCTACCGTTTGTGGTGTAGTAATCTAATACC 1113
DB 2996 TCCTGAAATAAATGTGGTGCCATCGAGGGGTCTACCGTTTGTGGTGTAGTAATCTAATACC 3055
QY 1114 TTAACCCCATCTCTCAAAATCAAAATGATCATATTTCTGAGAGACCCAGCAATACATA 1173
DB 3056 TTAACCCCATCTCTCAAAATCAAAATGATCATATTTCTGAGAGACCCAGCAATACATA 3115
QY 1174 AGAATTACTAAAAAATAAAAAA 1197
DB 3116 AGAATTACTAAAAAATAAAAAA 3139

RESULT 5

US-10-415-122-5
; Sequence 5, Application US/10415122
; Publication No. US20040053369A1
; GENERAL INFORMATION:
; APPLICANT: THE UNIVERSITY OF SYDNEY
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES
; FILE REFERENCE: FP15217
; CURRENT APPLICATION NUMBER: US/10/415,122
; CURRENT FILING DATE: 2003-08-07
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 3120
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-415-122-5

Query Match 86.5%; Score 1036; DB 7; Length 3120;
Best Local Similarity 89.0%; Pred. No. 1.4e-301;
Matches 1193; Conservative 0; Mismatches 0; Indels 147; Gaps 1;

QY 1 AATTTGAAGCCACCAAGACTCCCTTTAGAGCATCACCTGTAGTACGTACGTACGTAA 60
DB 1781 AATTTGAAGCCACCAAGACTCCCTTTAGAGCATCACCTGTAGTACGTACGTAA 1840
QY 61 ATCTGAGAGGTGACAGGCTGACGTACCGTGGCTACTCATCTTGTGTCATCAGTC 120
DB 1841 ATCTGAGAGGTGACAGGCTGACGTACCGTGGCTACTCATCTTGTGTCATCAGTC 1900
QY 121 AGCACTGTGACTCTTTATAGTAACTAGTAAACAGAGAAATCCACACTGTGTGCC 180
DB 1901 AGCACTGTGACTCTTTATAGTAACTAGTAAACAGAGAAATCCACACTGTGTGCC 1960
QY 181 TTTTACAGCTATCAAGTCTCTGAAGATGACCAACTTGCACAGAAATTTTGGGCCA 240
DB 1961 TTTTACAGCTATCAAGTCTCTGAAGATGACCAACTTGCACAGAAATTTTGGGCCA 2020
QY 241 CCAATTTTGGATTCAGAGGTCTCTTCTGACTATATCTCTCCAGAAATTTTCTTTTG 300
DB 2021 CCAATTTTGGATTCAGAGGTCTCTTCTGACTATATCTCTCCAGAAATTTTCTTTTG 2080
QY 301 AAAGTACTACTGGAATTTACATTTATGGAATGCTCTACAAGCTCATGATCTACAGCTG 360
DB 2081 AAAGTACTACTGGAATTTACATTTATGGAATGCTCTACAAGCTCATGATCTACAGCTG 2140
QY 361 GAAAGAAATATCTCTGCTGCTTTCATATATGTTGGTCTCA----- 403
DB 2141 GAAAGAAATATCTCTGCTGCTTTCATATATGTTGGTCTCAAGTGTGATGTAATA 2200
QY 404 ----- 403

RESULT 6

US-10-825-632-2
; Sequence 2, Application US/10825632
; Publication No. US20040191826A1
; GENERAL INFORMATION:
; APPLICANT: ABBOTT, Catherine Anne
; APPLICANT: GORRELL, Mark Douglas
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES
; FILE REFERENCE: FCSB-100-Div. 1

DB 2201 ATCCGTTTAAAGGAGTCAGTATTTCCTCGCTTGAATACCTAGCCTCTCTAGGTTATGTGG 2260
QY 404 ----- 403
DB 2261 TTCTAGTGATAGACAAACAGGGGATCTCTGTCACCGAGGGCTTAAATTTTGAAGGGCCCTTTA 2320
QY 404 -----GGGTCAAAATAGAAATTTGACGATCAGGTGGAAGGACTCCAAATATCTAGCTT 453
DB 2321 AATATAAATGGGTCAAAATAGAAATTTGACGATCAGGTGGAAGGACTCCAAATATCTAGCTT 2380
QY 454 CTCGATATGATTTTCAATTTGATCTTGTGGGATCCACGGTGTGCTTATGAGGAT 513
DB 2381 CTCGATATGATTTTCAATTTGATCTTGTGGGATCCACGGTGTGCTTATGAGGAT 2440
QY 514 ACCTCTCCCTGATGCGATTAATGACAGGTCAGATATCTTCAGGGTGTCTATTCCTCGGG 573
DB 2441 ACCTCTCCCTGATGCGATTAATGACAGGTCAGATATCTTCAGGGTGTCTATTCCTCGGG 2500
QY 574 CCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAACGTTATATGGTCCACC 633
DB 2501 CCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAACGTTATATGGTCCACC 2560
QY 634 CTGACCAGATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAGTTCC 693
DB 2561 CTGACCAGATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAGTTCC 2620
QY 694 CCTCTGAACCAATCGTTTACTGCTCTTACATGTTTCTCTGATGAGAAATGTCATTTTG 753
DB 2621 CCTCTGAACCAATCGTTTACTGCTCTTACATGTTTCTCTGATGAGAAATGTCATTTTG 2680
QY 754 CACATACCACTATTTACTGAGTTTTTGTAGGGCTGGAAGCCATATGATTTACAGA 813
DB 2681 CACATACCACTATTTACTGAGTTTTTGTAGGGCTGGAAGCCATATGATTTACAGA 2740
QY 814 TCTATCTCTAGGAGACACAGGATGAAGTTCTTGAATCGGAGAACATATGAACTGC 873
DB 2741 TCTATCTCTAGGAGACACAGGATGAAGTTCTTGAATCGGAGAACATATGAACTGC 2800
QY 874 ATCTTTTGGCACTACTCTTCAAGAAACCTTGGATCAGCTATTGCTCTCTAAAAGTGATAT 933
DB 2801 ATCTTTTGGCACTACTCTTCAAGAAACCTTGGATCAGCTATTGCTCTCTAAAAGTGATAT 2860
QY 934 AATTTTGACCTGTGTAGAACTCTCTGATATACACTGGCTATTTAACCAAAATGAGGAGTT 993
DB 2861 AATTTTGACCTGTGTAGAACTCTCTGATATACACTGGCTATTTAACCAAAATGAGGAGTT 2920
QY 994 TAATCAACAGAAACACAGAAATGATCATCATATTTGATACCTGCCATGTAAACATCTAC 1053
DB 2921 TAATCAACAGAAACACAGAAATGATCATCATATTTGATACCTGCCATGTAAACATCTAC 2980
QY 1054 TCCTGAAATAAATGTGGTGCCATGCGGGTCTACCGTTTGTGGTGTAGTAATCTAATACC 1113
DB 2981 TCCTGAAATAAATGTGGTGCCATGCGGGTCTACCGTTTGTGGTGTAGTAATCTAATACC 3040
QY 1114 TTAACCCCATCTCTCAAAATCAAAATGATCATATTTCTGAGAGACCCAGCAATACCAT 1173
DB 3041 TTAACCCCATCTCTCAAAATCAAAATGATCATATTTCTGAGAGACCCAGCAATACCAT 3100
QY 1174 AGAATTACTAAAAAATAAAAA 1193
DB 3101 AGAATTACTAAAAAATAAAAA 3120

```
; CURRENT APPLICATION NUMBER: US/10/825,632
; CURRENT FILING DATE: 2004-04-15
; PRIOR APPLICATION NUMBER: US 10/070,464
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: PCT/AU00/01085
; PRIOR FILING DATE: 2000-09-11
; PRIOR APPLICATION NUMBER: AU P05709
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: AU P02762
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH 3120
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-10-825-632-2

Query Match      86.5%; Score 1036; DB 8; Length 3120;
Best Local Similarity 89.0%; Pred. No. 1.4e-301;
Matches 1193; Conservative 0; Mismatches 0; Indels 147; Gaps 1;

QY 1 ATTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTAGCTAGTCAGTTAGGTAA 60
DB ATTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTAGCTAGTCAGTTAGGTAA 1840
QY 61 ATCTGGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACATTTCTGTGTCATCAGTC 120
DB ATCTGGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACATTTCTGTGTCATCAGTC 1900
QY 121 AGCACTGTGACTTCTTTTAAAGTAAAGTATAGTAACCAAGAGAAATCCACACTGTGTCC 180
DB AGCACTGTGACTTCTTTTAAAGTAAAGTATAGTAACCAAGAGAAATCCACACTGTGTCC 1960
QY 181 TTTTCAAGCTATCAAGTCTGGAAGATGACCAACTTGCACCAAGAGAAATTTTGGGCA 240
DB TTTTCAAGCTATCAAGTCTGGAAGATGACCAACTTGCACCAAGAGAAATTTTGGGCA 2020
QY 241 CCATTTTGGATTACAGAGTCTCTTCTGCTAGTATCTCTCCAGAAATTTCTCTTTG 300
DB CCATTTTGGATTACAGAGTCTCTTCTGCTAGTATCTCTCCAGAAATTTCTCTTTG 2080
QY 301 AAGTACTACTGGATTACATTTGATGGATGCTCTCAAGCCTCATGATCTACAGCCTG 360
DB AAGTACTACTGGATTACATTTGATGGATGCTCTCAAGCCTCATGATCTACAGCCTG 2140
QY 361 GAAAGAAATATCTTACTGTGCTGTTTCATATATGTTGTTCTCTCA----- 403
DB GAAAGAAATATCTTACTGTGCTGTTTCATATATGTTGTTCTCTCAAGGTCAGTTGGTGAATA 2200
QY 404 ----- 403
DB ATCGGTTTAAAGGATCAAGTATTTCCGCTTGAATACCCTAGCCTCTCTAGGTTATGTGG 2260
QY 404 ----- 403
DB TTGTAGTATAGACAACAGGGGATCCTGTGTACCGAGGGCTTAAATTTGAAGGCGCTTTTA 2320
QY 404 -----GGGTCAATAGAAATTTGACGATCAGGTGGAAGGACTCCAATATCTAGCTT 453
DB AATATAAAATGGGTCAATAGAAATTTGACGATCAGGTGGAAGGACTCCAATATCTAGCTT 2380
QY 454 CTCGATATGATTTTCAATTGACTTAGATCGTGTGGGCATCCACGGCTGGTCTCTATGAGGAT 513
DB CTCGATATGATTTTCAATTGACTTAGATCGTGTGGGCATCCACGGCTGGTCTCTATGAGGAT 2440
QY 514 ACCTCTCCCTGATGGCATTAAATGCAGAGGTGAGATATCTTTCAGGGTTGCTATTGCTGGG 573
DB ACCTCTCCCTGATGGCATTAAATGCAGAGGTGAGATATCTTTCAGGGTTGCTATTGCTGGG 2500
QY 574 CCCCACTGCTCTGTGGATCTTCTATGATACAGATACAGGAACGTTATATGGGTGACC 633
DB CCCCACTGCTCTGTGGATCTTCTATGATACAGATACAGGAACGTTATATGGGTGACC 2560
```

```
QY 634 CTGACCAAGATGAACAGGGGCTATTACTTAGGATCTGTGGCCATGCAACAGAAAGTTCC 693
DB CTGACCAAGATGAACAGGGGCTATTACTTAGGATCTGTGGCCATGCAACAGAAAGTTCC 2620
QY 694 CCTCTGAACCAAAATCGTTTACTGCTCTTACATGGTTTCTCTGATGAGAATGTCCATTTTG 753
DB CCTCTGAACCAAAATCGTTTACTGCTCTTACATGGTTTCTCTGATGAGAATGTCCATTTTG 2680
QY 754 CACATACCAGTATATTACTGAGTTTTTTAGTGAGGGCTGGAAAGCCATATGATTTACAGA 813
DB CACATACCAGTATATTACTGAGTTTTTTAGTGAGGGCTGGAAAGCCATATGATTTACAGA 2740
QY 814 TCTATCCTCAGGAGACACAGATAAGAGTTTCTGAAATCGGGAGAACATTATCAACTGC 873
DB TCTATCCTCAGGAGACACAGATAAGAGTTTCTGAAATCGGGAGAACATTATCAACTGC 2800
QY 874 ATCTTTTGCACTACCTTCAAGAAAACCTTGGATCACGTTATTGCTCTTAAAGTGATAT 933
DB ATCTTTTGCACTACCTTCAAGAAAACCTTGGATCACGTTATTGCTCTTAAAGTGATAT 2860
QY 934 AATTTTGACCTGTGTAGAACTCTCTGGTATACACTGGCTATTTAACCAAAATGAGGAGTT 993
DB AATTTTGACCTGTGTAGAACTCTCTGGTATACACTGGCTATTTAACCAAAATGAGGAGTT 2920
QY 994 TAATCAACAGAAACACAGAAATTGATCATCATTTTGTATACCTGCCATGTAACTCTAC 1053
DB TAATCAACAGAAACACAGAAATTGATCATCATTTTGTATACCTGCCATGTAACTCTAC 2980
QY 1054 TCCTGAAAATAAATGTGGTGCATCGAGGGTCTACGGTTTGGTGTAGTAACTCTAATACC 1113
DB TCCTGAAAATAAATGTGGTGCATCGAGGGTCTACGGTTTGGTGTAGTAACTCTAATACC 3040
QY 1114 TTAACCCGATGCTCAAAATCAATGATATATATTTCTTGAGAGACCAGCAATACCATA 1173
DB TTAACCCGATGCTCAAAATCAATGATATATATTTCTTGAGAGACCAGCAATACCATA 3100
QY 1174 AGAATTACTAAAAAATAAAA 1193
DB AGAATTACTAAAAAATAAAA 3120
```

```
RESULT 7
US-09-976-674-14
; Sequence 14, Application US/09976674
; Patent No. US20020115843A1
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
; LENGTH: 4309
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-976-674-14
```

```
Query Match      86.3%; Score 1033.2; DB 3; Length 4309;
Best Local Similarity 88.7%; Pred. No. 1.2e-300;
Matches 1194; Conservative 0; Mismatches 3; Indels 149; Gaps 1;

QY 1 ATTTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTAGCTAGTCAGTTAGGTAA 60
DB ATTTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTAGCTAGTCAGTTAGGTAA 1320
```


QY 403 ----- 402
Db 1681 TAATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCCCTAGCCTCTCTAGTTATGT 1740
QY 403 ----- 402
Db 1741 GGTGTAGTATAGACAACAGGGGATCTGTCAACGAGGGCTTAAATTTGAAGCGCCTT 1800
QY 403 -----AGGTCAAATAGAAATTGACGATCAGGTGGAAGGACTCCCAATATCTAGC 451
Db 1801 TAAATATAAATGGGTCAATAGAAATTGACGATCAGGTGGAAGGACTCCCAATATCTAGC 1860
QY 452 TTCTCGATATGATTTCAATTGACTTAGATCGTGTGGGCATCCACGGCTGGTCTCTATGGAGG 511
Db 1861 TTCTCGATATGATTTCAATTGACTTAGATCGTGTGGGCATCCACGGCTGGTCTCTATGGAGG 1920
QY 512 ATACCTCTCCCTGATGSCATTAATGACAGAGGTGAGATATCTTCAGGGTGTGATTTGGTGG 571
Db 1921 ATACCTCTCCCTGATGSCATTAATGACAGAGGTGAGATATCTTCAGGGTGTGATTTGGTGG 1980
QY 572 GGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAACTTATATGGGTCA 631
Db 1981 GGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAACTTATATGGGTCA 2040
QY 632 CCCTGACCAAGTGAACAGGGCTATTACTTAGATCTGTGGCCATGCAAGCAGAAAGTT 691
Db 2041 CCCTGACCAAGTGAACAGGGCTATTACTTAGATCTGTGGCCATGCAAGCAGAAAGTT 2100
QY 692 CCCTCTGAACCAATCGTTTACTGCTCTTACATGGTTTCTGGATGAGATGTCATTT 751
Db 2101 CCCTCTGAACCAATCGTTTACTGCTCTTACATGGTTTCTGGATGAGATGTCATTT 2160
QY 752 TGCATATACAGTATATTTACTGAGTTTTTTAGTAGGGCTGGAAAGCCATATGATTTACA 811
Db 2161 TGCATATACAGTATATTTACTGAGTTTTTTAGTAGGGCTGGAAAGCCATATGATTTACA 2220
QY 812 GATCTATCTCAGGAGAGACACAGCATAAAGATTCCTGAATCGGAGAACATTATGAAT 871
Db 2221 GATCTATCTCAGGAGAGACACAGCATAAAGATTCCTGAATCGGAGAACATTATGAAT 2280
QY 872 GCATTTTTGCACTACTCTTCAAGAAACCTTGGATCAGTATTTGCTCTTAAAGTGTAT 931
Db 2281 GCATTTTTGCACTACTCTTCAAGAAACCTTGGATCAGTATTTGCTCTTAAAGTGTAT 2340
QY 932 ATATTTTGGACCTGTGTAGAACTCTCTGGTATACACTGGCTATTTTAAACCAATGAGGAGG 991
Db 2341 ATATTTTGGACCTGTGTAGAACTCTCTGGTATACACTGGCTATTTTAAACCAATGAGGAGG 2400
QY 992 TTTAATCAACAGAAAAACAGAAATTGATCATCACATTTTGTATACCTGCAATGTAAACATCT 1051
Db 2401 TTTAATCAACAGAAAAACAGAAATTGATCATCACATTTTGTATACCTGCAATGTAAACATCT 2460
QY 1052 ACTCTGAAATATAATGTGTGCCATGCAAGGGTCTACGGTTTGTGTGTAGTAAATCTAATA 1111
Db 2461 ACTCTGAAATATAATGTGTGCCATGCAAGGGTCTACGGTTTGTGTGTAGTAAATCTAATA 2520
QY 1112 CCTTAACCCACATGCTCAAAATCAATGATACATATTCCTGAGAGAGCCAGCAATACCA 1171
Db 2521 CCTTAACCCACATGCTCAAAATCAATGATACATATTCCTGAGAGAGCCAGCAATACCA 2580
QY 1172 TAAGATTACTTAAAAAANAAAAA 1197
Db 2581 TAAGATTACTTAAAAAANAAAAA 2606

RESULT 9
US-09-976-674-12
; Sequence 12, Application US/09976674
; Patent No. US20020115943A1
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre

; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 4829
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-976-674-12

Query Match 86.3%; Score 1033.2; DB 3; Length 4829;
Best Local Similarity 88.7%; Pred. No. 1-2e-300;
Matches 1194; Conservative 0; Mismatches 3; Indels 149; Gaps 1;

QY 1 ATTTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTACGTAGTACGTACGTAA 60
Db 1781 ATTTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTACGTAGTACGTAA 1840
QY 61 ATCTGAGAGGTGACCAAGCTGACTGACCGTGGCTACTCACATTTCTGCTGCATCAGTC 120
Db 1841 ATCTGAGAGGTGACCAAGCTGACTGACCGTGGCTACTCACATTTCTGCTGCATCAGTC 1900
QY 121 AGCACTGTGACTTCTTTTAAAGTATAGTAAACAGAAATCCACATCTGTGTGTCCT 180
Db 1901 AGCACTGTGACTTCTTTTAAAGTATAGTAAACAGAAATCCACATCTGTGTGTCCT 1960
QY 181 TTTTCAAGCTATCAAGTCTGAGATGACCCAACTTGCAGAAACAAAGAAATTTTGGGCCA 240
Db 1961 TTTTCAAGCTATCAAGTCTGAGATGACCCAACTTGCAGAAACAAAGAAATTTTGGGCCA 2020
QY 241 CCATTTTGGATTTCAGCAGGTCCTTCTCTGACTATACCTCTCCAGAAATTTTCTCTTTTG 300
Db 2021 CCATTTTGGATTTCAGCAGGTCCTTCTCTGACTATACCTCTCCAGAAATTTTCTCTTTTG 2080
QY 301 AAAGTACTACTGGAATTTACATTTGTATGGGATGCTCTACAAGCTCATGATCTACAGCCTG 360
Db 2081 AAAGTACTACTGGAATTTACATTTGTATGGGATGCTCTACAAGCTCATGATCTACAGCCTG 2140
QY 361 GAAAGAAATATCCTACTGTGCTGTTTCAATATATGGTGGTCTCTCCTCAGGTGCAAGTGTG 402
Db 2141 GAAAGAAATATCCTACTGTGCTGTTTCAATATATGGTGGTCTCTCCTCAGGTGCAAGTGTG 2200
QY 403 ----- 402
Db 2201 TAATCGGTTTAAAGAGTCAAGTATTTTCGGCTTGAATACCCCTAGCCTCTCTAGGTTATGT 2260
QY 403 ----- 402
Db 2261 GGTGTAGTATAGACAACAGGGGATCCTGTCAACGAGGGCTTAAATTTGAAGCGCCTT 2320
QY 403 -----AGGTCAAATAGAAATTGACGATCAGGTGGAAGGACTCCCAATATCTAGC 451
Db 2321 TAAATATAAATGGGTCAATAGAAATTGACGATCAGGTGGAAGGACTCCCAATATCTAGC 2380
QY 452 TTCTCGATATGATTTCAATTGACTTAGATCGTGTGGGCATCCACGGCTGGTCTCTATGGAGG 511
Db 2381 TTCTCGATATGATTTCAATTGACTTAGATCGTGTGGGCATCCACGGCTGGTCTCTATGGAGG 2440
QY 512 ATACCTCTCCCTGATGSCATTAATGACAGAGGTGAGATATCTTCAGGGTGTGATTTGGTGG 571
Db 2441 ATACCTCTCCCTGATGSCATTAATGACAGAGGTGAGATATCTTCAGGGTGTGATTTGGTGG 2500
QY 572 GGCCCCAGTCACTCTCTGTGGATCTTCTATGATACAGGATACACGGAACTTATATGGGTCA 631
Db 2501 GGCCCCAGTCACTCTCTGTGGATCTTCTATGATACAGGATACACGGAACTTATATGGGTCA 2560
QY 632 CCCTGACCAAGTGAACAGGGCTATTACTTAGGATCTCTGTGGCCATGCAAGCAGAAAGTT 691

Db	2561	CGCTGACCAGATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAAAGTT	2620	Db	1841	ATCCTGAGAGGTGACAAGGCTGACTGAACCGTGGCTACTCACATTTCTTGCTGCATCAGTC	1900
Qy	692	CCCTCTGAACCAATCGTTTACTGCTCTTACATGGTTTCTCGATGAGATGCCATTT	751	Qy	121	AGCACTGTGACTTCTTTTATAAGTAAGTATAGTACCAGAAATCCACACTGTGTGCC	180
Db	2621	CCCTCTGAACCAATCGTTTACTGCTCTTACATGGTTTCTCGATGAGATGCCATTT	2680	Db	1901	AGCACTGTGACTTCTTTTATAAGTAAGTATAGTACCAGAAATCCACACTGTGTGCC	1960
Qy	752	TGCACATACCAAGTATATTAAGTCTGAGTTTCTGAGGGCTGGAAGCCATATGATTTACA	811	Qy	181	TTTACAAAGCTATCAAGTCTGAAAGATGACCCAACTTGCACAAACAAAGAAATTTTGGGCCA	240
Db	2681	TGCACATACCAAGTATATTAAGTCTGAGTTTCTGAGGGCTGGAAGCCATATGATTTACA	2740	Db	1961	TTTACAAAGCTATCAAGTCTGAAAGATGACCCAACTTGCACAAACAAAGAAATTTTGGGCCA	2020
Qy	812	GATCTATCCTCAGGAGACACAGATGAAGAGTTCTGAAATCGGGAGAACATTAAGAACT	871	Qy	241	CCATTTTGGATTTCAGCAGGTCTCTTCTCTGATATATCTCTCCAGAAATTTTCTTTTG	300
Db	2741	GATCTATCCTCAGGAGACACAGATGAAGAGTTCTGAAATCGGGAGAACATTAAGAACT	2800	Db	2021	CCATTTTGGATTTCAGCAGGTCTCTTCTCTGATATATCTCTCCAGAAATTTTCTTTTG	2080
Qy	872	GCATCTTTTGCATACCTTTCAAGAAACCTTTGGATCAGTATTCCTCTTAAAGTGAT	931	Qy	301	AAAGTACTACTGATTTACATTTGATGGATGCTCTCAAGCCTCATGATCTACAGCTG	360
Db	2801	GCATCTTTTGCATACCTTTCAAGAAACCTTTGGATCAGTATTCCTCTTAAAGTGAT	2860	Db	2081	AAAGTACTACTGATTTACATTTGATGGATGCTCTCAAGCCTCATGATCTACAGCTG	2140
Qy	932	ATAATTTTGACCTGTGAGAACTCTCTGGTATACACTGGCTATTTAAACCAATGAGGAG	991	Qy	361	GAAGAAATATCTTACTGCTGCTTCATATATATGTTGGTCTCTCTCAGGTGACAGTTGGTGAA	402
Db	2861	ATAATTTTGACCTGTGAGAACTCTCTGGTATACACTGGCTATTTAAACCAATGAGGAG	2920	Db	2141	GAAGAAATATCTTACTGCTGCTTCATATATATGTTGGTCTCTCTCAGGTGACAGTTGGTGAA	2200
Qy	992	TTTAATCAACAGAAACACAGAAATGATCATCACAATTTTGATCCTGCAATGAAATCT	1051	Qy	403	-----	402
Db	2921	TTTAATCAACAGAAACACAGAAATGATCATCACAATTTTGATCCTGCAATGAAATCT	2980	Db	2201	TAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCCCTAGGCTCTCTAGGTTATGT	2260
Qy	1052	ACTCCTGAAATTAATGCTGCCATGAGGGGCTACGGTTTGTGGTAGTAATCTAATA	1111	Qy	403	-----	402
Db	2981	ACTCCTGAAATTAATGCTGCCATGAGGGGCTACGGTTTGTGGTAGTAATCTAATA	3040	Db	2261	GGTTGTAGTATAGACAAACAGGGGATCTGTCCACCGGGCTTAAATTTGAAGGCGCTT	2320
Qy	1112	CCTTAACCCCAATGCTCAAAATCAAAATGATACATATTTCTGAGAGACCCAGCAATACCA	1171	Qy	403	-----	451
Db	3041	CCTTAACCCCAATGCTCAAAATCAAAATGATACATATTTCTGAGAGACCCAGCAATACCA	3100	Db	2321	TAAATATAAAATGGGTCAAATAGAAATTCAGATCAGGTGGAAGACTCCAAATATCTAGC	2380
Qy	1172	TAAAGATTAATAAAAAAAAAAAAAA	1197	Qy	452	TTCTCGATATGATTTCAATTTGATGATGCTGTGGGCATCCACGGCTGGTCTCTATGAGG	511
Db	3101	TAAAGATTAATAAAAAAAAAAAAAA	3126	Db	2381	TTCTCGATATGATTTCAATTTGATGATGCTGTGGGCATCCACGGCTGGTCTCTATGAGG	2440
RESULT 10							
US-10-982-512-12							
; Sequence 12, Application US/10982512							
; Publication No. US20050059081A1							
; GENERAL INFORMATION:							
; APPLICANT: Qi, Steve							
; APPLICANT: Akinsanya, Karen							
; APPLICANT: Riviere, Pierre							
; APPLICANT: Junien, Jean-Louis							
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV							
; FILE REFERENCE: 70669							
; CURRENT APPLICATION NUMBER: US/10/982,512							
; CURRENT FILING DATE: 2004-11-05							
; PRIOR APPLICATION NUMBER: US/09/976,674							
; PRIOR FILING DATE: 2001-10-12							
; PRIOR APPLICATION NUMBER: US 60/240,117							
; PRIOR FILING DATE: 2000-10-12							
; NUMBER OF SEQ ID NOS: 61							
; SOFTWARE: PatentIn version 3.1							
; SEQ ID NO 12							
; LENGTH: 4829							
; TYPE: DNA							
; ORGANISM: Homo sapiens							
US-10-982-512-12							
Query Match 86.3%; Score 1033.2; DB 9; Length 4829;							
Best Local Similarity 88.7%; Pred. No. 1.2e-300;							
Matches 1194; Conservative 0; Mismatches 3; Indels 149; Gaps 1;							
Qy	1	ATTTTGAAGGACCAAGAGCTCCCTTTAGAGATCAGTCTGACCTAGCTAGTACCTGTA	60	Qy	992	TTTAATCAACAGAAACACAGAAATGATCATCACAATTTTGATCCTGCAATGAAATCT	1051
Db	1781	ATTTTGAAGGACCAAGAGCTCCCTTTAGAGATCAGTCTGACCTAGCTAGTACCTGTA	1840	Db	2921	TTTAATCAACAGAAACACAGAAATGATCATCACAATTTTGATCCTGCAATGAAATCT	2980
Qy	61	ATCCTGAGAGGTGACAAAGGCTGACTGACCGTGGCTACTCACATTTCTTGCTGCATCAGTC	120	Qy	403	-----	402


```
Db 2921 TTTAATCAACAGAAAACACAGAAATTGATCATCACATTTTGTATACCTGCCATGTAACATCT 2980
Qy 1052 ACTCCTGAAATAAATGTGGTGCATGCGAGGGGTCTACGGTTTGTGGTGTAGTAATCTTAATA 1111
Db 2981 ACTCCTGAAATAAATGTGGTGCATGCGAGGGGTCTACGGTTTGTGGTGTAGTAATCTTAATA 3040
Qy 1112 CTTTAAACCCACATGCTCAAAATCAAAATGATACATATTCCTGAGAGACCCAGCAATACCA 1171
Db 3041 CTTTAAACCCACATGCTCAAAATCAAAATGATACATATTCCTGAGAGACCCAGCAATACCA 3100
Qy 1172 TAAGAAATTACTAAAAAARAAAAA 1197
Db 3101 TAAGAAATTACTAAAAAARAAAAA 3126

RESULT 11
US-10-311-035-30
; Sequence 30: Application US/10311035
; Publication No. US20040023243A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: YUE, Henry
; APPLICANT: ELLIOTT, Vicki
; APPLICANT: GANDHI, Ameena R.
; APPLICANT: LAL, Preeti
; APPLICANT: AU-YOUNG, Janice
; APPLICANT: TRIBOULEY, Catherine M.
; APPLICANT: DELEGEANE, Angelo M.
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: NGUYEN, Dannie B.
; APPLICANT: LEE, Ernestine A.
; APPLICANT: HAFALIA, April
; APPLICANT: KHAN, Farrah A.
; APPLICANT: CHAWLA, Narinder K.
; APPLICANT: YAO, Monique G.
; APPLICANT: LU, Dying Aina M.
; APPLICANT: ARVIZU, Chandra S.
; APPLICANT: TANG, Y. Tom
; APPLICANT: WALSH, Roderick T.
; APPLICANT: AZIMZAI, Valda
; APPLICANT: LU, Yan
; APPLICANT: RAMKUMAR, Jayalaximi
; APPLICANT: XU, Yuming
; APPLICANT: REDDY, Roopa
; APPLICANT: DAS, Depodriya
; APPLICANT: KEARNEY, Liam
; APPLICANT: KALLICK, Deborah A.
; TITLE OF INVENTION: proteases
; FILE REFERENCE: FI-0123 PCT
; CURRENT APPLICATION NUMBER: US/10/311,035
; PRIOR FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: 60/212,336; 60/213,955; 60/215,396; 60/216,821; 60/218,946
; PRIOR FILING DATE: 2000-06-16; 2000-06-22; 2000-06-29; 2000-07-07; 2000-07-14
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PERL Program
; SEQ ID NO 30
; LENGTH: 3106
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. US20040023243A1 7160544CB1
US-10-311-035-30

Query Match 86.3%; Score 1033; DB 7; Length 3106;
Best Local Similarity 89.0%; Pred. No. 1.1e-300;
Matches 1190; Conservative 0; Mismatches 0; Indels 147; Gaps 1;

Qy 1 ATTTTGAAGGCCAACAAAGACTCCCTTTAGAGCATCACCTGTACGTAGTCAGTTACGTAA 60
Db 1770 ATTTTGAAGGCCAACAAAGACTCCCTTTAGAGCATCACCTGTACGTAGTCAGTTACGTAA 1829
Qy 61 ATCCTGGAGAGGTGACAAAGGCTGACTGACCGTGCTACTCACATTTCTTCTGCTGCATCAGTC 120
```

```
Db 1830 ATCCTGGAGAGGTGACAAAGGCTGACTGACCGTGGCTACTCACTTTCTTGTGCATCAGTC 1889
Qy 121 AGCACTGTGACTTCTTTTATAGTAAGTATAGTAACACAGAAATCCACACTGTGTGTCCC 180
Db 1890 AGCACTGTGACTTCTTTTATAGTAAGTATAGTAACACAGAAATCCACACTGTGTGTCCC 1949
Qy 181 TTTTCAAGCTATCAAGTCTCTGAAGATGACCCAACTTGC AAAA CAAAGAAATTTTGGGCCA 240
Db 1950 TTTTCAAGCTATCAAGTCTCTGAAGATGACCCAACTTGC AAAA CAAAGAAATTTTGGGCCA 2009
Qy 241 CCATTTTGGATTACAGCAGGTCTCTTCTGTGACTATATCTCCTCCAGAAAATTTTCTCTTTG 300
Db 2010 CCATTTTGGATTACAGCAGGTCTCTTCTGTGACTATATCTCCTCCAGAAAATTTTCTCTTTG 2069
Qy 301 AAAGTACTACTGGATTTCATTTGTATGGGATGCTCTACAAGCCTCATGATCTACAGCCTG 360
Db 2070 AAAGTACTACTGGATTTCATTTGTATGGGATGCTCTACAAGCCTCATGATCTACAGCCTG 2129
Qy 361 GAAAGAAATATCCTACTGTGCTGTTTATATATGTTGGTGTCTCA----- 403
Db 2130 GAAAGAAATATCCTACTGTGCTGTTTATATATGTTGGTGTCTCAGGTGCAGTTGTTGAATA 2189
Qy 404 ----- 403
Db 2190 ATCGGTTTAAAGGTCAGATATTTCCGCTTGAATACCTAGACCTCTCTAGTTATCTGG 2249
Qy 404 ----- 403
Db 2250 TTGTAGTGATAGACAACAGGGGATCCTGTCCACCGAGGGCTTAAATTTGAAGGCGCCTTTA 2309
Qy 404 -----GGGTCAATAGAAATAGAGATCAGGTGAGGAGGACTCCAAATATCTAGCTT 453
Db 2310 AATATAAATGGGTCAATAGAAATGACATCAGGTGGAAGGACTCAATATCTAGCTT 2369
Qy 454 CTCGATATGATTTTCAATGATCTAGATCGTGTGGGATCCACGGCTGTCTCTATGGAGAT 513
Db 2370 CTCGATATGATTTTCAATGATCTAGATCGTGTGGGATCCACGGCTGTCTCTATGGAGAT 2429
Qy 514 ACCTCTCCCTGATGGCATTAATGACAGGTGAGATATCTTCAGGGTGTCTATTTGCTGGG 573
Db 2430 ACCTCTCCCTGATGGCATTAATGACAGGTGAGATATCTTCAGGGTGTCTATTTGCTGGG 2489
Qy 574 CCCAGTCACTCTGTGGATCTTTATGATACAGGATACACGGAAACGTTATATGGTCAAC 633
Db 2490 CCCAGTCACTCTGTGGATCTTTATGATACAGGATACACGGAAACGTTATATGGTCAAC 2549
Qy 634 CTGACCAGAAATGAACAGGGCTATTACTTAGGATCTGTGGCATCTGAAAGCAAGAAAGTTCC 693
Db 2550 CTGACCAGAAATGAACAGGGCTATTACTTAGGATCTGTGGCATCTGAAAGCAAGAAAGTTCC 2609
Qy 694 CCTCTGAACCAATCGTTTACTGTCTTACATGTTTCTCTGGATGAGAAATGTCCATTTTG 753
Db 2610 CCTCTGAACCAATCGTTTACTGTCTTACATGTTTCTCTGGATGAGAAATGTCCATTTTG 2669
Qy 754 CACATACCAAGTATATTACTGAGTTTTTTAGTGGGCTGGAAAGCCATATGATTTACAGA 813
Db 2670 CACATACCAAGTATATTACTGAGTTTTTTAGTGGGCTGGAAAGCCATATGATTTACAGA 2729
Qy 814 TCTATCCTCAGGAGAGACACAGATAGAGTTCTGTAATCGGAGACATTTATGAACTGC 873
Db 2730 TCTATCCTCAGGAGAGACACAGATAGAGTTCTGTAATCGGAGACATTTATGAACTGC 2789
Qy 874 ATCTTTTGCACTACCTTCAAGAAACCTTGGATCACGTTATTTGCTGTAAAGATGAT 933
Db 2790 ATCTTTTGCACTACCTTCAAGAAACCTTGGATCACGTTATTTGCTGTAAAGATGAT 2849
Qy 934 AATTTTGACCTGTGTAGAACTCTCTGGTATACACTGGCTATTAAACAAATGAGGAGTT 993
Db 2850 AATTTTGACCTGTGTAGAACTCTCTGGTATACACTGGCTATTAAACAAATGAGGAGTT 2909
Qy 994 TAAATCAAGAAAACACAGAAATTCATCACATTTTGTATACCTGCCATGTAACTCTAC 1053
```



```
Db 2910 TAAATCAACAGAAACACAGAAATTGATCATCATCAATTTTGATACCTGCCATGTAAACATCTAC 2969
Qy 1054 TCTGAAATAAATGTGGTCCCATGACGGGGTCTACGGTTTGTGGTAGTAATCTAATACC 1113
Db 2970 TCTGAAATAAATGTGGTCCCATGACGGGGTCTACGGTTTGTGGTAGTAATCTAATACC 3029
Qy 1114 TTAACCCACATCTCAAAATCAAAATGATACATATTTCTGAGAGACCCAGCAATACCATA 1173
Db 3030 TTAACCCACATCTCAAAATCAAAATGATACATATTTCTGAGAGACCCAGCAATACCATA 3089
Qy 1174 AGAATTACTAAAAAAA 1190
Db 3090 AGAATTACTAAAAAAA 3106

RESULT 12
US-09-976-674-22
; Sequence 22, Application US/09976674
; Patent No. US20020115843A1
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 22
; LENGTH: 4685
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-976-674-22

Query Match 84.7%; Score 1013.6; DB 3; Length 4685;
Best Local Similarity 91.3%; Pred. No. 1e-294;
Matches 1098; Conservative 0; Mismatches 99; Indels 5; Gaps 2;

Qy 1 ATTTTGAAGCACCAGAACTCCCTTTAGAGCATCACTGTAGTAGTCAAGTTACGTAA 60
Db 1781 ATTTTGAAGCACCAGAACTCCCTTTAGAGCATCACTGTAGTAGTCAAGTTACGTAA 1840
Qy 61 ATCTGTGAGAGGTGACAAAGCTGACTGACCGTGGCTACTCAATTTCTTGTGCAATCAGTC 120
Db 1841 ATCTGTGAGAGGTGACAAAGCTGACTGACCGTGGCTACTCAATTTCTTGTGCAATCAGTC 1900
Qy 121 AdCACTGTGACTTCTTTTATAAGTAAGTATAGTAACAGAGAAATCCACACTGTGTGCC 180
Db 1901 AdCACTGTGACTTCTTTTATAAGTAAGTATAGTAACAGAGAAATCCACACTGTGTGCC 1960
Qy 181 TTTCACAGCTATCAAGTCTGAAAGTACCCAACTTGCACAAAGAAATTTTGGGCCA 240
Db 1961 TTTCACAGCTATCAAGTCTGAAAGTACCCAACTTGCACAAAGAAATTTTGGGCCA 2020
Qy 241 CCATTTGGATTGAC-----CAGGTCTCTTCTGTGACT-ATACTCTCCAGAAATTTTCTC 295
Db 2021 CCATTTGGATTGACCTCCTCAGGTGCGATTGGTGAATAATCGGTTTAAAGAGTCAAGTA 2080
Qy 296 TTTTGAAGTACTACTGGAATTTACATTGATGGATGCTCTACAGCCTCATGATCTACA 355
Db 2081 TTTCCGCTTGAATACCTCAGCTCTCTAGGTATGTTGGTTGATGATAGACACAGGGG 2140
Qy 356 GCCTGGAAGAAATATCCTACTGTCTGTATATATGTTGCTCTCAGGGGTCAATAGA 415
Db 2141 ATCTGTGACCGAGGGCTTAAATTTGAGCGCCCTTAAATATAAATGGGTCAATAGA 2200
Qy 416 AATTGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATGATGTTTCATGACTT 475
Db 2201 AATTGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATGATGTTTCATGACTT 2260
```

```
Qy 476 AGATCGTGTGGGCATCCACGGCTGTGTCCTATGAGGATACCTCTCCCTGATGGCAATTAAT 535
Db 2261 AGATCGTGTGGGCATCCACGGCTGTGTCCTATGAGGATACCTCTCCCTGATGGCAATTAAT 2320
Qy 536 GCAGAGGTCAAGATATCTTTCAGGGTTGCTATTGCTGGGGCCCCAGTCACTCTGTGGATCTT 595
Db 2321 GCAGAGGTCAAGATATCTTTCAGGGTTGCTATTGCTGGGGCCCCAGTCACTCTGTGGATCTT 2380
Qy 596 CTATGATACAGGATACACGGACGTTATATGGGTCAACCTGACCCAGATGACAGGGCTA 655
Db 2381 CTATGATACAGGATACACGGACGTTATATGGGTCAACCTGACCCAGATGACAGGGCTA 2440
Qy 656 TTACTTAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTGAACCAAAATCGTTTACT 715
Db 2441 TTACTTAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTGAACCAAAATCGTTTACT 2500
Qy 716 GCTCTTACATGTTTCTCTGGATGAGAAATGTCATTTTGCACATACCATGATATTAATCTGAG 775
Db 2501 GCTCTTACATGTTTCTCTGGATGAGAAATGTCATTTTGCACATACCATGATATTAATCTGAG 2560
Qy 776 TTTTATAGTGAGGGCTGGAAAGCCATATGATTTACAGATCTATCTCAGGAGACACAG 835
Db 2561 TTTTATAGTGAGGGCTGGAAAGCCATATGATTTACAGATCTATCTCAGGAGACACAG 2620
Qy 836 CATAAGAGTTCCTGAATCGGAGAACATTTATGAATCTATCTGCACTACCTTTCAAGA 895
Db 2621 CATAAGAGTTCCTGAATCGGAGAACATTTATGAATCTATCTGCACTACCTTTCAAGA 2680
Qy 896 AAACCTTGGATCAGGTATTTGCTGCTCTAAAGTGATATAATTTTACCTGTGTAGAACTC 955
Db 2681 AAACCTTGGATCAGGTATTTGCTGCTCTAAAGTGATATAATTTTACCTGTGTAGAACTC 2740
Qy 956 TCTGGTATACACTGGCTATTTTAAACCAATGAGGAGTTTAAATCAACAGAAAACACAGAAAT 1015
Db 2741 TCTGGTATACACTGGCTATTTTAAACCAATGAGGAGTTTAAATCAACAGAAAACACAGAAAT 2800
Qy 1016 TGATCATCAATTTTGATACCTGCCATGTAAACATCTACTCTCGAAAATAAATGTGTGCC 1075
Db 2801 TGATCATCAATTTTGATACCTGCCATGTAAACATCTACTCTCGAAAATAAATGTGTGCC 2860
Qy 1076 ATGAGGGGTCTACGGTTTGTGTAGTAATCTAAATACCTTAACCCACATGCTCAAAATC 1135
Db 2861 ATGAGGGGTCTACGGTTTGTGTAGTAATCTAAATACCTTAACCCACATGCTCAAAATC 2920
Qy 1136 AAATGATACATATTTCTGAGAGACCCAGGAATACCAATAGATTAATTAATAAAAAA 1195
Db 2921 AAATGATACATATTTCTGAGAGACCCAGGAATACCAATAGATTAATTAATAAAAAA 2980
Qy 1196 AA 1197
Db 2981 AA 2982
```

```
RESULT 13
US-10-982-512-22
; Sequence 22, Application US/10982512
; Publication No. US20050059081A1
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/10/982,512
; CURRENT FILING DATE: 2004-11-05
; PRIOR APPLICATION NUMBER: US/09/976,674
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
```

```
; SEQ ID NO 22
; LENGTH: 4685
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-982-512-22

Query Match      84.7%; Score 1013.6; DB 9; Length 4685;
Best Local Similarity 91.3%; Pred. No. 1e-294;
Matches 1098; Conservative 0; Mismatches 99; Indels 5; Gaps 2;

Qy 1 ATTTGAGGACCAAGAAGTCTCCCTTTAGAGCATCACTGTAGCTAGTACGTAA 60
Db ATTTGAGGACCAAGAAGTCTCCCTTTAGAGCATCACTGTAGCTAGTACGTAA 1840
Qy 61 ATCTGGAGAGGTGACAAGGCTGACTGACCGGTGCTACTCAATCTTGTGTCATCAGTC 120
Db ATCTGGAGAGGTGACAAGGCTGACTGACCGGTGCTACTCAATCTTGTGTCATCAGTC 1900
Qy 121 AGCACTGTGACTCTTTTATAAGTAAGTATAGTAACCAAGAAATCCACACTGTGTGCC 180
Db AGCACTGTGACTCTTTTATAAGTAAGTATAGTAACCAAGAAATCCACACTGTGTGCC 1960
Qy 181 TTTACAAGCTATCAAGTCTCTGAAGATGACCCAACTTGCAAAACAAAGAAATTTTGGGCCA 240
Db TTTACAAGCTATCAAGTCTCTGAAGATGACCCAACTTGCAAAACAAAGAAATTTTGGGCCA 2020
Qy 241 CCATTTTGGATTTCAG---CAGTCTCTTCTGACT-ATACTCTCCAGAAATTTTCTC 295
Db CCATTTTGGATTTCAGTCTCTGAGTGGTGAATTAATCGGTTTAAAGAGTCAAGTA 2080
Qy 296 TTTTGAAGTACTACTGGATTACATTTGTATGGGATGCTCTACAAGCCTCATGATCTACA 355
Db TTTCCGCTTGAATACCTAGCTCTCTAGGTATAGTGGTGTAGTATAGACAACAGGGG 2140
Qy 356 GCTCGAAGAAATATCTCTGCTGCTGTTCATATATGGTGGTCTCAGGGTCAATATAGA 415
Db ATCTGTCCAGGCGCTTAAATTTGAAGGCGCTTTAAATATATAAATGGGTCAATATAGA 2200
Qy 416 AATTGACGATCAGTGGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTCACTT 475
Db AATTGACGATCAGTGGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTCACTT 2260
Qy 476 AGATCTGTGGGATCCAGCGGTGCTCTATGGAGGATACCTCTCCCTGATGGCATTAT 535
Db AGATCTGTGGGATCCAGCGGTGCTCTATGGAGGATACCTCTCCCTGATGGCATTAT 2320
Qy 536 GCAGAGGTGAGATATCTTACGGTTTGTCTATTTGCTGGGGCCCCAGTCACTCTGTGATCTT 595
Db GCAGAGGTGAGATATCTTACGGTTTGTCTATTTGCTGGGGCCCCAGTCACTCTGTGATCTT 2380
Qy 596 CTATGATACAGGATACAGGAACGTTATATGGGTCACTGACCAAGTGAACAGGCGCTA 655
Db CTATGATACAGGATACAGGAACGTTATATGGGTCACTGACCAAGTGAACAGGCGCTA 2440
Qy 656 TTACTTTAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCTCTGAAACCAATCGTTTACT 715
Db TTACTTTAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCTCTGAAACCAATCGTTTACT 2500
Qy 716 GCTCTTACATGTTTCTTGGATGAGATGTCATTTTGCATACCAAGTATATTTACTGAG 775
Db GCTCTTACATGTTTCTTGGATGAGATGTCATTTTGCATACCAAGTATATTTACTGAG 2560
Qy 776 TTTTGTAGTGGGCTGGAAGCCATATGATTTACATCTATTCCTCAGAGAGACAG 835
Db TTTTGTAGTGGGCTGGAAGCCATATGATTTACATCTATTCCTCAGAGAGACAG 2620
Qy 836 CATAGAGTCTCTGATCGGAGAACATATGAACTGATCTTTTGCATCACTCTTCAAGA 895
Db CATAGAGTCTCTGATCGGAGAACATATGAACTGATCTTTTGCATCACTCTTCAAGA 2680
Qy 896 AAACCTTGGATCAGCTATGCTGCTCTAAAAGTATATAAATTTGACCTGTGTGAACCTC 955
Db AAACCTTGGATCAGCTATGCTGCTCTAAAAGTATATAAATTTGACCTGTGTGAACCTC 2740
```

```
Qy 956 TCTGGTATACACTGGGCTATTTAACTCAAATGAGGAGGTTTAACTCAACAGAAAACACAGAA 1015
Db 2741 TCTGGTATACACTGGGCTATTTAACTCAAATGAGGAGGTTTAACTCAACAGAAAACACAGAA 2800
Qy 1016 TGATCATCACATTTTGTAGTACCTGCCATGTAACATCTACTCTCTGAAAAATAAATGTGGTGCC 1075
Db 2801 TGATCATCACATTTTGTAGTACCTGCCATGTAACATCTACTCTCTGAAAAATAAATGTGGTGCC 2860
Qy 1076 ATGAGGGGTCTACGGTTTGTGTAGTAACTTAATACCTTAAACCCCAATGCTCAAAATC 1135
Db 2861 ATGAGGGGTCTACGGTTTGTGTAGTAACTTAATACCTTAAACCCCAATGCTCAAAATC 2920
Qy 1136 AAATGATACATATCTCTGAGAGACCCAGCAATACCATTAAGAAATTACTAAAAAATAAATAA 1195
Db 2921 AAATGATACATATCTCTGAGAGACCCAGCAATACCATTAAGAAATTACTAAAAAATAAATAA 2980
Qy 1196 AA 1197
Db 2981 AA 2982

RESULT 14
US-10-825-632-6
; Sequence 6; Application US/10825632
; Publication No. US20040191826A1
; GENERAL INFORMATION:
; APPLICANT: ABSOTI; Catherine Anne
; APPLICANT: GORRELL, Mark Douglas
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES
; FILE REFERENCE: FCSB-100-Div. 1
; CURRENT APPLICATION NUMBER: US/10/825,632
; CURRENT FILING DATE: 2004-04-15
; PRIOR APPLICATION NUMBER: US 10/070,464
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: PCT/AU00/01085
; PRIOR FILING DATE: 2000-09-11
; PRIOR APPLICATION NUMBER: AU PQ5709
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: AU PQ2762
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1669
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-10-825-632-6

Query Match      73.6%; Score 881; DB 8; Length 1669;
Best Local Similarity 87.2%; Pred. No. 7.1e-255;
Matches 1044; Conservative 0; Mismatches 0; Indels 153; Gaps 1;

Qy 1 ATTTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTAGCTAGTACGTAGCTAA 60
Db 618 ATTTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTAGCTAGTACGTAGCTAA 677
Qy 61 ATCTGGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACATTTCTGTGATCAGTC 120
Db 678 ATCTGGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACATTTCTGTGATCAGTC 737
Qy 121 AGCATGTGACTCTTTTATAAGTATAGTATAGTAACCAAGAAATCCACACTGTGTGCC 180
Db 738 AGCATGTGACTCTTTTATAAGTATAGTATAGTAACCAAGAAATCCACACTGTGTGCC 797
Qy 181 TTTTACAAGCTATCAAGTCTCTGAAAGATGACCCAACTTGCAAAACAAAGAAATTTTGGGCCA 240
Db 798 TTTTACAAGCTATCAAGTCTCTGAAAGATGACCCAACTTGCAAAACAAAGAAATTTTGGGCCA 857
Qy 241 CCATTTTGGATTCAGCAGGCTCTCTTCTGACTATATCTCTCAGAAATTTTCTCTTTTG 300
Db 858 CCATTTTGGATTCAGCAGGCTCTCTTCTGACTATATCTCTCAGAAATTTTCTCTTTTG 917
```

301 AAAGTACTACTGGATTTACATTTGATGGAGTCTTACAAGCCTCATGATCTACAGCCTG 360
Db |||||
918 AAAGTACTACTGGATTTACATTTGATGGAGTCTTACAAGCCTCATGATCTACAGCCTG 977
Qy |||||
361 GAAAGAAATATCCTACTGCTGCTTTCATATATGTTGGTCTTCAAGGGTCAAAATAGAAATG 420
Db |||||
978 GAAAGAAATATCCTACTGCTGCTTTCATATATGTTGGTCTTCAAGGGTCTTCAAGGGTCTTCA 1020
Qy |||||
421 AAGATCAGGTGGAGAGGAGTCAATATCTAGCTTCTCGATATGATTTCAATGACTTAGATC 480
Db |||||
1021 ||||| 1020
Qy 481 GTGTGGGCATCCACGGCTGCTCTATGGAGGATACCTCTCCCTGATGCAATTAATGCAGA 540
Db ||||| 1021 ||||| 1020
Qy 541 GGTGAGATATCTTCAAGGTTGCTATTGCTGGGGCCCGAGTCACTCTGTGGATCTTCTATG 600
Db ||||| 1021 ||||| 1064
Qy 601 ATACAGGATACAGGAACGTTATATGAGTCACTTACCCCTGACCAAGTGAACAGGGCTATTACT 660
Db ||||| 1065 ATACAGGATACAGGAACGTTATATGAGTCACTTACCCCTGACCAAGTGAACAGGGCTATTACT 1124
Qy 661 TAGGATCTGTGGCCATGCAAGCAGAAAGTTCCCTCTGAACCAATCGTTTACTGCTCT 720
Db ||||| 1125 TAGGATCTGTGGCCATGCAAGCAGAAAGTTCCCTCTGAACCAATCGTTTACTGCTCT 1184
Qy 721 TACATGTTCTCTGGATGAGAAATGTCATTTTGCACATACCAAGTATATTACTAGTATTTT 780
Db ||||| 1185 TACATGTTCTCTGGATGAGAAATGTCATTTTGCACATACCAAGTATATTACTAGTATTTT 1244
Qy 781 TAGTGAGGCTGGAGAGCAATGATTTACAGATCTATCCCTCAGAGAGACACAGCATAA 840
Db ||||| 1245 TAGTGAGGCTGGAGAGCAATGATTTACAGATCTATCCCTCAGAGAGACACAGCATAA 1304
Qy 841 GAGTTCCTGAATCGGAGAACATTTAAGTCACTCTTTCGACCTTACCTTCAAGAAACC 900
Db ||||| 1305 GAGTTCCTGAATCGGAGAACATTTAAGTCACTCTTTCGACCTTACCTTCAAGAAACC 1364
Qy 901 TTGATCACGTAATGCTCTTAAAGTGAATATTTTGACCTGTGTAGAACCTCTCTGG 960
Db ||||| 1365 TTGATCACGTAATGCTCTTAAAGTGAATATTTTGACCTGTGTAGAACCTCTCTGG 1424
Qy 961 TATACATGCTATTTAAACCAATGAGAGGTTTAAATCAACAGAAACACAGAAATGATC 1020
Db ||||| 1425 TATACATGCTATTTAAACCAATGAGAGGTTTAAATCAACAGAAACACAGAAATGATC 1484
Qy 1021 ATACATTTTGATACCTGCAATGATCACTCTTCAACCAATGAGAGGTTTAAATCAACAGAAACACAGAAATGATC 1080
Db ||||| 1485 ATACATTTTGATACCTGCAATGATCACTCTTCAACCAATGAGAGGTTTAAATCAACAGAAACACAGAAATGATC 1544
Qy 1081 GGGGTCTACGTTTGTGATGATTAATCTAATACCTTAAACCAATGAGAGGTTTAAATCAACAGAAACACAGAAATGATC 1140
Db ||||| 1545 GGGGTCTACGTTTGTGATGATTAATCTAATACCTTAAACCAATGAGAGGTTTAAATCAACAGAAACACAGAAATGATC 1604
Qy 1141 ATACATATTTCTGAGAGACCCAGCAATACCAATGATTAATCAACAGAAACACAGAAATGATC 1197
Db ||||| 1605 ATACATATTTCTGAGAGACCCAGCAATACCAATGATTAATCAACAGAAACACAGAAATGATC 1661

RESULT 15
US-10-956-157-2177
; Sequence 2177, Application US/10956157
; Publication No. US20050118625A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts William
; TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
; FILE OF INVENTION: HUMAN OSTEOARTHRITIS AND HUMAN PROTEASES
; FILE REFERENCE: 031896-043000 (AM 101081)
; CURRENT APPLICATION NUMBER: US/10/956,157
; CURRENT FILING DATE: 2004-10-04

; NUMBER OF SEQ ID NOS: 319805
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2177
; LENGTH: 2830
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-956-157-2177

Query Match 73.6%; Score 881; DB 9; Length 2830;
Best Local Similarity 87.2%; Pred. No; 9.7e-255;
Matches 1044; Conservative 0; Mismatches 0; Indels 153; Gaps 1;

Qy 1 ATTTTGAAGGACCAAGAGCTCCCTTTAGAGCATCACCTGTACGTAGTACGTACGTA 60
Db ||||| 1781 ATTTTGAAGGACCAAGAGCTCCCTTTAGAGCATCACCTGTACGTAGTACGTACGTA 1840
Qy 61 ATCTGTGAGAGGTGACAAAGGCTGACTGAGCGTGGGTACTCATTCTTGTGTCATCAGTC 120
Db ||||| 1841 ATCTGTGAGAGGTGACAAAGGCTGACTGAGCGTGGGTACTCATTCTTGTGTCATCAGTC 1900
Qy 121 AGCACTGTGACTTCTTTTATAAGTAAGTATAGTAACCAAGAAATCCACACTGTGTGTC 180
Db ||||| 1901 AGCACTGTGACTTCTTTTATAAGTAAGTATAGTAACCAAGAAATCCACACTGTGTGTC 1960
Qy 191 TTTTCAAGCTATCAAGTCTGAAGATGACCCAACTTGCACAAAGAAATTTTGGGCCA 240
Db ||||| 1961 TTTTCAAGCTATCAAGTCTGAAGATGACCCAACTTGCACAAAGAAATTTTGGGCCA 2020
Qy 241 CCATTTTGGATTCAGCAGGTCTCTTCTGACTATCTCTCCAGAAATTTCTCTTTTG 300
Db ||||| 2021 CCATTTTGGATTCAGCAGGTCTCTTCTGACTATCTCTCCAGAAATTTCTCTTTTG 2080
Qy 301 AAAGTACTACTGGATTTACATTTGATGGAGTCTTACAAGCCTCATGATCTACAGCCTG 360
Db ||||| 2081 AAAGTACTACTGGATTTACATTTGATGGAGTCTTACAAGCCTCATGATCTACAGCCTG 2140
Qy 361 GAAAGAAATATCTTACTGCTCTTCAATATATGTTGGTCTCTCAGGGTCAAAATAGAAATG 420
Db ||||| 2141 GAAAGAAATATCTTACTGCTCTTCAATATATGTTGGTCTCTCAGGGTCAAAATAGAAATG 2183
Qy 421 ACATCAGGTGGAAGGACTCCAAATATCTAGCTTCTCGATATGATTTCAATGACTTAGATC 480
Db ||||| 2184 ||||| 2183
Qy 481 GTGTGGGCATCCACGGCTGCTCTATGGAGGATACCTCTCCCTGATGCAATTAATGCAGA 540
Db ||||| 2184 ||||| 2183
Qy 541 GGTGAGATATCTTCAAGGTTGCTATTGCTGGGGCCCGAGTCACTCTGTGGATCTTCTATG 600
Db ||||| 2184 ||||| 2227
Qy 601 ATACAGGATACAGGAACGTTATATGAGTCACTCTTCAACCAATGAGAGGCTATTACT 660
Db ||||| 2228 ATACAGGATACAGGAACGTTATATGAGTCACTCTTCAACCAATGAGAGGCTATTACT 2287
Qy 661 TAGGATCTGTGGCCATGCAAGCAGAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCT 720
Db ||||| 2288 TAGGATCTGTGGCCATGCAAGCAGAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCT 2347
Qy 721 TACATGTTCTCTGGATGAGAAATGTCATTTTGCACATACCAAGTATATTACTAGTATTTT 780
Db ||||| 2348 TACATGTTCTCTGGATGAGAAATGTCATTTTGCACATACCAAGTATATTACTAGTATTTT 2407
Qy 781 TAGTGAGGCTGGAAAGCCATATGATTTACAGATCTATCTCTCAGAGAGACACAGCATAA 840
Db ||||| 2408 TAGTGAGGCTGGAAAGCCATATGATTTACAGATCTATCTCTCAGAGAGACACAGCATAA 2467
Qy 841 GAGTTCCTGAATCGGAGAACATTTAAGTCACTCTTTCGACCTTACCTTCAAGAAACC 900
Db ||||| 2468 GAGTTCCTGAATCGGAGAACATTTAAGTCACTCTTTCGACCTTACCTTCAAGAAACC 2527
Qy 901 TTGATCACGTAATGCTCTTAAAGTGAATATTTTGACCTGTGTAGAACCTCTCTGG 960

Db	2528	TTGGATCAGGTATTGGCTGCTTAAAGTGTATATAATTTTGACCTGTGTGAACTCTCTGG	2587
Qy	961	TATACACTGGCTATTTAAACCAATGAGGAGGTTTAATCAACAGAAAAACACAGAATTGATC	1020
Db	2588	TATACACTGGCTATTTAAACCAATGAGGAGGTTTAATCAACAGAAAAACACAGAATTGATC	2647
Qy	1021	ATCACATTTTGATACCTGGCCATGTAAACATCTACTCTGAAAAATAAATGTGTGCCATGCA	1080
Db	2648	ATCACATTTTGATACCTGGCCATGTAAACATCTACTCTGAAAAATAAATGTGTGCCATGCA	2707
Qy	1081	GGGTCTACGGTTTGTGGTAGTAACTAATACCTTTAACCCACATGCTCAAAATCAAATG	1140
Db	2708	GGGTCTACGGTTTGTGGTAGTAACTAATACCTTTAACCCACATGCTCAAAATCAAATG	2767
Qy	1141	ATACATATTCCTGAGAGACCCAGCAATACCATAGAATTACTAAAAA	1197
Db	2768	ATACATATTCCTGAGAGACCCAGCAATACCATAGAATTACTAAAAA	2824

Search completed: May 4, 2006, 11:21:58
Job time : 1070.84 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 4, 2006, 02:53:29 ; Search time 5966.21 Seconds
(without alignments)
11404.501 Million cell updates/sec

Title: US-10-825-632-4
Perfect score: 1197
Sequence: 1 atttgaagcaccacaagac.....ttactaaaaaaaaaaaaa 1197

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues
Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*
1: gb_ba.*
2: gb_in.*
3: gb_env.*
4: gb_on.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pr.*
9: gb_ro.*
10: gb_sts.*
11: gb_sy.*
12: gb_un.*
13: gb_vi.*
14: gb_htg.*
15: gb_pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1197	100.0	1197	6 AR651453	AR651453 Sequence
2	1197	100.0	1197	8 AF221635	AF221635 Homo sapi
3	1179	98.5	4676	6 AR631287	AR631287 Sequence
4	1179	98.5	4676	6 AX608743	AX608743 Sequence
5	1040	86.9	2797	6 AR448400	AR448400 Sequence
6	1040	86.9	3127	8 AF221634	AF221634 Homo sapi
7	1040	86.9	3143	6 AX354793	AX354793 Sequence
8	1036	86.5	3120	6 AR651452	AR651452 Sequence
9	1036	86.5	4535	8 BC040203	BC040203 Homo sapi
10	1033.2	86.3	4309	6 AR631284	AR631284 Sequence
11	1033.2	86.3	4309	6 AX608737	AX608737 Sequence
12	1033.2	86.3	4829	6 AR631283	AR631283 Sequence
13	1033.2	86.3	4829	6 AX608735	AX608735 Sequence
14	1033	86.3	3106	6 AX342633	AX342633 Sequence
15	1016.2	84.9	3125	8 BC030688	BC030688 Homo sapi
16	1013.6	84.7	4685	6 AR631288	AR631288 Sequence
17	1013.6	84.7	4685	6 AX608745	AX608745 Sequence
18	999.6	83.5	2161	6 BD157001	BD157001 Primer fo

19	999.6	83.5	2161	6	AX878058	AX878058 Sequence
20	999.6	83.5	2161	8	AK027826	AK027826 Homo sapi
21	923	77.1	3030	8	AY354202	AY354202 Homo sapi
22	881	73.6	1669	6	AR651454	AR651454 Sequence
23	881	73.6	1669	8	AF221636	AF221636 Homo sapi
24	881	73.6	2778	8	AK000290	AK000290 Homo sapi
25	879	73.4	2668	6	AX405771	AX405771 Sequence
26	879	73.4	2842	6	AX405770	AX405770 Sequence
27	873	72.9	4523	6	AR631281	AR631281 Sequence
28	873	72.9	4523	6	AX608731	AX608731 Sequence
29	863.6	72.1	2510	6	AX338497	AX338497 Sequence
30	791.4	66.1	873	6	AX083130	AX083130 Sequence
31	778	65.0	2649	6	AX354795	AX354795 Sequence
32	776.4	64.9	2656	8	AF172659	AF172659 Homo sapi
33	776.4	64.9	2671	6	AR631278	AR631278 Sequence
34	776.4	64.9	2671	6	AX608725	AX608725 Sequence
35	704.8	58.9	4799	9	BC043124	BC043124 Mus muscu
36	704.8	58.9	4799	9	BC059222	BC059222 Mus muscu
37	677	56.6	735	6	AX524942	AX524942 Sequence
38	633.4	52.9	823	6	AR263993	AR263993 Sequence
39	599.2	50.1	1002	6	AX524939	AX524939 Sequence
40	540	45.1	2098	6	CQ721272	CQ721272 Sequence
41	523.6	43.7	587	6	BD154822	BD154822 Primer fo
42	523.6	43.7	587	6	AX874760	AX874760 Sequence
43	500.6	41.8	1449	5	CR353491	CR353491 Gallus ga
44	459	38.3	1083	6	AR651455	AR651455 Sequence
45	459	38.3	1083	8	AF221637	AF221637 Homo sapi

ALIGNMENTS

RESULT 1
AR651453
LOCUS AR651453 1197 bp DNA linear PAT 20-APR-2005
DEFINITION Sequence 4 from patent US 6881564.
ACCESSION AR651453
VERSION AR651453.1 GI:62795939
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1197)
AUTHORS Abbott,C.A. and Gorrell,M.D.
TITLE Dipeptidyl peptidases
JOURNAL Patent: US 6881564-A 4 19-Apr-2005;
The University of Sydney; Sydney;
AUX;
FEATURES
source Location/Qualifiers
1..1197
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN	Query Match	Best Local Similarity	Score	DB 6;	Length	1197;
	Matches 1197;	Conservative	0;	Mismatches	0;	Indels
						Gaps
						0;
Qy	1	ATTTTGAAGGCACCAAGACTCCCTTTTAGACATCACCTGTACGTAGTCAGTACGTAA	60			
Db	1	ATTTTGAAGGCACCAAGACTCCCTTTTAGACATCACCTGTACGTAGTCAGTACGTAA	60			
Qy	61	ATCCTGGAGAGGTGACCAAGGCTGACTGACCGTGGCTACTCACATTTCTGTCATCAGTC	120			
Db	61	ATCCTGGAGAGGTGACCAAGGCTGACTGACCGTGGCTACTCACATTTCTGTCATCAGTC	120			
Qy	121	AGCACTGTGACTTCTTTATAGTAAGTATAGTAACCAAGAAATCCACACTGTGTGTGCC	180			
Db	121	AGCACTGTGACTTCTTTATAGTAAGTATAGTAACCAAGAAATCCACACTGTGTGTGCC	180			
Qy	181	TTTACAGCTATCAAGTCTTCAAGTCCCTGAAAGTACCCAACTTGCAGAAAGAAATTTTGGCCA	240			
Db	181	TTTACAGCTATCAAGTCTTCAAGTCCCTGAAAGTACCCAACTTGCAGAAAGAAATTTTGGCCA	240			

Qy	301	AAAGTACTACTGGATTACATTTGATGGGATGCTCTACAAGCCTCATGATCTACAGCCTG	360	TITLE	Serine protease genes related to DPPIV			
Db	301	AAAGTACTACTGGATTACATTTGATGGGATGCTCTACAAGCCTCATGATCTACAGCCTG	360	JOURNAL	Patent: US 6844180-A 20 18-JAN-2005;			
Qy	361	GAAGAATAATCTACTGTGCTGTTTCATATATGGTGGTCTCTCAGGGTCAAAATGAATAATG	420		Ferring BV;;			
Db	361	GAAGAATAATCTACTGTGCTGTTTCATATATGGTGGTCTCTCAGGGTCAAAATGAATAATG	420	FEATURES	NLX;			
Qy	421	ACGATCAGGTGGAAGGACTCCAAATATCTAGCTTCTCGATATGATTTCTTACGATCTAGATC	480	source	Location/Qualifiers			
Db	421	ACGATCAGGTGGAAGGACTCCAAATATCTAGCTTCTCGATATGATTTCTTACGATCTAGATC	480		1. 4676			
Qy	481	GTGTGGGCATCCACGGCTGTCTCTATGAGGATACCTCTCCCTGATGSCATTATATGAGA	540	ORIGIN	/organism="unknown"			
Db	481	GTGTGGGCATCCACGGCTGTCTCTATGAGGATACCTCTCCCTGATGSCATTATATGAGA	540		/mol_type="genomic DNA"			
Qy	541	GGTCAGATATCTTTCAGGGTTGCTATTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATG	600	Query Match	98.5%; Score 1179; DB 6; Length 4676;			
Db	541	GGTCAGATATCTTTCAGGGTTGCTATTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATG	600	Best Local Similarity	99.7%; Pred. No. 6.7e-304;			
Qy	601	ATACAGGATACACGGAAAGTATATGGGTCAACCTGACAGAAATGAACAGGGCTATTACT	660	Matches 1193; Conservative	0; Mismatches 0; Indels 4; Gaps 1;			
Db	601	ATACAGGATACACGGAAAGTATATGGGTCAACCTGACAGAAATGAACAGGGCTATTACT	660	Qy	1	ATTTTGAAGSCACCAAGACTCCCTTTTAGAGCATCACCTGTACGTAGTCAGTTACGTAA	60	
Qy	661	TAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTTGAAACCAAAATCGTTTACTGCTCT	720	Db	1781	ATTTTGAAGSCACCAAGACTCCCTTTTAGAGCATCACCTGTACGTAGTCAGTTACGTAA	1840	
Db	661	TAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTTGAAACCAAAATCGTTTACTGCTCT	720	Qy	61	ATCCTGTGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACAATTTCTGTGTCATCAGTC	120	
Qy	721	TACATGGTTTCTCGGATGAGAAATGTCATTTTGCAATACACAGTATATTACTGAGTTTTT	780	Db	1841	ATCCTGTGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACAATTTCTGTGTCATCAGTC	1900	
Db	721	TACATGGTTTCTCGGATGAGAAATGTCATTTTGCAATACACAGTATATTACTGAGTTTTT	780	Qy	121	AGCACTGTGACTTCTTTTATAAGTAAGTATAGTAAACAGAGAATCCACACTGTGTGTCCC	180	
Qy	781	TAGTGGGCTGGAAAGCCATATGATTTACAGATCTATCTTCAGAGAGACACAGCATAA	840	Db	1901	AGCACTGTGACTTCTTTTATAAGTAAGTATAGTAAACAGAGAATCCACACTGTGTGTCCC	1960	
Db	781	TAGTGGGCTGGAAAGCCATATGATTTACAGATCTATCTTCAGAGAGACACAGCATAA	840	Qy	181	TTTACAAGCTATCAAGTCTCTGAAGATGACCCAACTTTGCAAAAACAAGGAATTTTGGGCCA	240	
Qy	841	GAGTTCCTGAAATCGGAGAACATTATGAATCTGATCTTTTGCACTTACCTTCAAGAAAACC	900	Db	1961	TTTACAAGCTATCAAGTCTCTGAAGATGACCCAACTTTGCAAAAACAAGGAATTTTGGGCCA	2020	
Db	841	GAGTTCCTGAAATCGGAGAACATTATGAATCTGATCTTTTGCACTTACCTTCAAGAAAACC	900	Qy	241	CCATTTTGGATTCAGCAGAGTCTCTTCTGACTATATCTCTCCACGAAATTTTCTCTTTTG	300	
Qy	901	TTGGATCACGTATTTGCTCTCTAAAGTGTATATAATTTTGACCTGTGTAGAACTCTCTGG	960	Db	2021	CCATTTTGGATTCAGCAGAGTCTCTTCTGACTATATCTCTCCACGAAATTTTCTCTTTTG	2080	
Db	901	TTGGATCACGTATTTGCTCTCTAAAGTGTATATAATTTTGACCTGTGTAGAACTCTCTGG	960	Qy	301	AAAGTACTACTGGAATTTACATTTGATGGGATGCTCTACAAGCCTCATGATCTACAGCCTG	360	
Qy	961	TATACACTGGCTATTTAAACAAATGAGGAGTTTAAATCAACAGAAAACACAGAAATTCATC	1020	Db	2081	AAAGTACTACTGGAATTTACATTTGATGGGATGCTCTACAAGCCTCATGATCTACAGCCTG	2140	
Db	961	TATACACTGGCTATTTAAACAAATGAGGAGTTTAAATCAACAGAAAACACAGAAATTCATC	1020	Qy	361	GAAGAATAATCTACTGTGCTGTCATATATGTTGGTCTCTCAGGGTCAAAATAGAAAATG	420	
Qy	1021	ATCACATTTTGATACCTGCCATGTAAATCTACTCTCGAAATATAATGTTGGTCCATGCA	1080	Db	2141	GAAGAATAATCTACTGTGCTGTCATATATGTTGGTCTCTCAGGGTCAAAATAGAAAATG	2196	
Db	1021	ATCACATTTTGATACCTGCCATGTAAATCTACTCTCGAAATATAATGTTGGTCCATGCA	1080	Qy	421	ACGATCAGGTGGAGGAGTCCAAATATCTAGCTTCTCGATATGATTTCTAGCTTAGATC	480	
Qy	1081	GAGGCTACGGTTTGTGTAGTAACTTAATACCTTAACCCACATGCTCAAAATCAAAATG	1140	Db	2197	ACGATCAGGTGGAGGAGTCCAAATATCTAGCTTCTCGATATGATTTCTAGCTTAGATC	2256	
Db	1081	GAGGCTACGGTTTGTGTAGTAACTTAATACCTTAACCCACATGCTCAAAATCAAAATG	1140	Qy	481	GTGTGGGCATCCACGGCTGCTCTATGAGGATACCTCTCCCTGATGCAATTAATGCGAGA	540	
Qy	1141	ATACATATTTCTGAGAGACCCAGCAATACATAAGAAATTTACTTAAAAAAA	1197	Db	2257	GTGTGGGCATCCACGGCTGCTCTATGAGGATACCTCTCCCTGATGCAATTAATGCGAGA	2316	
Db	1141	ATACATATTTCTGAGAGACCCAGCAATACATAAGAAATTTACTTAAAAAAA	1197	Qy	541	GGTCAGATATCTTTCAGGGTTGCTATTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATG	600	
RESULT 3	AR631287	Sequence 20 from patent US 6844180.	linear	Db	2317	GGTCAGATATCTTTCAGGGTTGCTATTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATG	2376	
LOCUS	AR631287	4676 bp	DNA	Qy	601	ATACAGGATACACGGAAAGTATATGGGTCAACCTGATGSCATTATATGAGATTAAT	660	
DEFINITION	Sequence 20 from patent US 6844180.			Db	2377	ATACAGGATACACGGAAAGTATATGGGTCAACCTGATGSCATTATATGAGATTAAT	2436	
ACCESSION	AR631287			Qy	661	TAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTGAAACCAAAATCGTTTACTGCTCT	720	
VERSION	AR631287.1	GI:59770934		Db	2437	TAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTGAAACCAAAATCGTTTACTGCTCT	2496	
KEYWORDS	Unknown.			Qy	721	TACATGGTTTCTCGGATGAGAAATGTCATTTTGCAATACACAGTATATTACTGAGTTTTT	780	
SOURCE	Unknown.			Db	2497	TACATGGTTTCTCGGATGAGAAATGTCATTTTGCAATACACAGTATATTACTGAGTTTTT	2556	
ORGANISM	Unclassified.			Qy	781	TAGTGGGCTGGAAAGCCATATGATTTACAGATCTATCTCTCAGGAGAGACACAGCATAA	840	
REFERENCE	1 (bases 1 to 4676)			Db	2557	TAGTGGGCTGGAAAGCCATATGATTTACAGATCTATCTCTCAGGAGAGACACAGCATAA	2616	
AUTHORS	Oi, S., Akinsanya, K.O., Riviere, P.J.M. and Junien, J.-L.			Qy	841	GAGTTCCTGAAATCGGAGAACATTATGAATCTGATCTTTTGTGCACTACCTTCAAGAAAACC	900	
				Db	2617	GAGTTCCTGAAATCGGAGAACATTATGAATCTGATCTTTTGTGCACTACCTTCAAGAAAACC	2676	

/db_xref="GI:11095188"
LKKLADTRKXVHGMKAPLDFPDKNDPDGPHSDRIYVILAMSGENRNTLFTISEI
PKTINRAAIVLHVMKPLDLFOALDQYMYSEBELAERXIGTIGASIDYHQSG
TELFQAGSGIYHVKGDPGPGFTQPLBNLJUTSPNRMDFKUCPADPWIAPHSN
DIWISNIVTFRPRLJYVNELANNEEDARAGUNATVYQBEFDRISGYWMCPEKETT
PSGGKILRILYENDSESEVEIITHVSEVLESTRADSFRIKGTIGANPKVTFKMSIEMI
DARGRLIDVDELLOPRILFEQVEYANAGWTFEGYIAWSILLDRSQRLQIVLIS
PELPIVEDDVMERGLIESVEDSVPLIITETIDMINIHDIHFVFPQSHBEIEF
IPASECKTFRHLXYKITSLEKSKYRSRSGGULPAPDFKCIKEEIAITSGEWEVLGR
HGSNIQVDRVHLVPEGTDSFLEHLIYVSNYPGEVTKLDRGYSHCSCISQBCD
FETLSYGNKPNCHSLYKLSKSPEDPTCKTKEFWATILDSAGPLPYTPPEIFSES
TGTFTYGCYKPHDLQPKKPYTFLFYGGQVOLVNNRPGVKYPRLNTLASGYV
VVUVDNRGSCYKPHDLQPKKPYTFLFYGGQVOLVNNRPGVKYPRLNTLASGYV
GVYLSMALMQSDIFRVAIAGAPVILWIFDYDTGYRMYGHPDQEQYLGVSAMQ
AEXFPEPRLHLLGFDLENVHFAFTSILLFLVRAGKPYDLQIYVQPERHSIRVPES
GEHYELHLHYLOENLJGSRIAAKVI"

ORIGIN

Query Match	86.9%;	Score 1040;	DB 8;	Length 3127;
Best Local Similarity	89.1%;	Pred. No. 9.7e-267;		
Matches 1197;	Conservative 0;	Mismatches 0;	Indels 147;	Gaps 1;
Qy	1	ATTTGAGGACCAAGAAGTCCCTTTAGAGCATCACCTGACGTAGTCAGTTAGCTAA	60	
Db	1781	ATTTGAGGACCAAGAAGTCCCTTTAGAGCATCACCTGACGTAGTCAGTTAGCTAA	1840	
Qy	61	ATCCTGAGAGGTGCAAGAGGTGACTGACCGTGGTACTCACATTTCTTGCTGCATCAGTC	120	
Db	1841	ATCCTGAGAGGTGCAAGAGGTGACTGACCGTGGTACTCACATTTCTTGCTGCATCAGTC	1900	
Qy	121	AGCACTGTGACTCTTTTAAAGTAAAGTATAGTACCAAGAGAAATCCACACTGTGTCCC	180	
Db	1901	AGCACTGTGACTCTTTTAAAGTAAAGTATAGTACCAAGAGAAATCCACACTGTGTCCC	1960	
Qy	181	TTTACAAGCTATCAAGTCTGAAGATGACCAACTTGCACAAACAAAGGAATTTGGGCCA	240	
Db	1961	TTTACAAGCTATCAAGTCTGAAGATGACCAACTTGCACAAACAAAGGAATTTGGGCCA	2020	
Qy	241	CCATTTTGGATTACAGAGTCTCTTCTGCTGACTATCTCTCCAGAAATTTCTCTTTG	300	
Db	2021	CCATTTTGGATTACAGAGTCTCTTCTGCTGACTATCTCTCCAGAAATTTCTCTTTG	2080	
Qy	301	AAAGTACTACTGATTTTACATTTGATGGATGCTCTACRAGCCTCATGATCTACAGCCTG	360	
Db	2081	AAAGTACTACTGATTTTACATTTGATGGATGCTCTACRAGCCTCATGATCTACAGCCTG	2140	
Qy	361	GAAGAATAATCTACTGCTGTTCATATATGTTGCTCTCA	403	
Db	2141	GAAGAATAATCTACTGCTGTTCATATATGTTGCTCTCA	2200	
Qy	404	-----	403	
Db	2201	ATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCCCTAGCCTCTCTAGGTTATGTGG	2260	
Qy	404	-----	403	
Db	2261	TTGTAGTATAGACACACAGGGATCTGTGCCAGGGCTTAAATTTGAAGCGCCTTTA	2320	
Qy	404	-----GGGTCAATAGAAATGACGATCAGGTGGAAGACTCCAAATATCTAGCTT	453	
Db	2321	AATATAAATGGGTCAATAGAAATGACGATCAGGTGGAAGACTCCAAATATCTAGCTT	2380	
Qy	454	CTCGATATGATTTCAATTGACTTAGATCGTGTGGGCATCCACGGCTGGTCTCTATGAGGAT	513	
Db	2381	CTCGATATGATTTCAATTGACTTAGATCGTGTGGGCATCCACGGCTGGTCTCTATGAGGAT	2440	
Qy	514	ACCTCTCCCTGATGGCAATTAATGACAGAGGTGAGATATCTTTCAGGGTGTGATTTGCTGGG	573	
Db	2441	ACCTCTCCCTGATGGCAATTAATGACAGAGGTGAGATATCTTTCAGGGTGTGATTTGCTGGG	2500	
Qy	574	CCCCAGTCACCTCTGTGGATCTTCTATGATACAGATACACGAAAGTTATATGGGTACCC	633	

Db	2501	CCCCAGTCACCTCTGTGGATCTTCTATGATACAGGATACACGGAAGCTTATATGGGTACCC	2560	
Qy	634	CTGACCAAGATGAACAGGGCTATTACTTAGGATCTCTGGCCCATGCAAGCAGAAAAATTTCC	693	
Db	2561	CTGACCAAGATGAACAGGGCTATTACTTAGGATCTCTGGCCCATGCAAGCAGAAAAATTTCC	2620	
Qy	694	CCTCTGAACCAAAATCGTTTACTCTCTTACATGTTTCTGATGAGAAATGTCCTATTTTG	753	
Db	2621	CCTCTGAACCAAAATCGTTTACTCTCTTACATGTTTCTGATGAGAAATGTCCTATTTTG	2680	
Qy	754	CACATACCACTATATTACTAGTATTTTGTAGTCAGGCTGGAAGCCATATGATTTACAGA	813	
Db	2681	CACATACCACTATATTACTAGTATTTTGTAGTCAGGCTGGAAGCCATATGATTTACAGA	2740	
Qy	814	TCTATCCTCAGGAGAGACACAGCATAAGAGTTTCTGAAATCGGGAGAACATTATGAACTGC	873	
Db	2741	TCTATCCTCAGGAGAGACACAGCATAAGAGTTTCTGAAATCGGGAGAACATTATGAACTGC	2800	
Qy	874	ATCTTTTGGCACTACCTTTCAAGAAACCTTGGATCAGTATTGCTCTTAAAGTGATAT	933	
Db	2801	ATCTTTTGGCACTACCTTTCAAGAAACCTTGGATCAGTATTGCTCTTAAAGTGATAT	2860	
Qy	934	AATTTTGACCTGTAGAACTCTCTGGTATACACTGCTGCTATTTAACCAATGAGAGGTT	993	
Db	2861	AATTTTGACCTGTAGAACTCTCTGGTATACACTGCTGCTATTTAACCAATGAGAGGTT	2920	
Qy	994	TAATCAACAGAAACACAGAAATGATCATCATTTTGTATACCTGCCATGTAACATCTAC	1053	
Db	2921	TAATCAACAGAAACACAGAAATGATCATCATTTTGTATACCTGCCATGTAACATCTAC	2980	
Qy	1054	TCTGAAATTAATGTGGTGCCATCGAGGGGTCTACGGTTTGTGGTAGTAACTTAATACC	1113	
Db	2981	TCTGAAATTAATGTGGTGCCATCGAGGGGTCTACGGTTTGTGGTAGTAACTTAATACC	3040	
Qy	1114	TTAACCCCATGCTCTCAAAATCAAAATGATACATTTCTCTGAGAGACCAGCAATACCATA	1173	
Db	3041	TTAACCCCATGCTCTCAAAATCAAAATGATACATTTCTCTGAGAGACCAGCAATACCATA	3100	
Qy	1174	AGAAATTAATCAAAAAAATAAAAAA 1197		
Db	3101	AGAAATTAATCAAAAAAATAAAAAA 3124		

RESULT 7
AX354793
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
CDS

AX354793
Sequence 1 from Patent WO0179473.
AX354793
AX354793.1 GI:18619526
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
1
Meyers, R.A. and Williamson, M.
21953, a human prolyl oligopeptidase family member and uses thereof
Patent: WO 0179473-A 1 25-OCT-2001.
Millennium Pharmaceuticals, Inc. (US)
Location/Qualifiers
1. 3143
/organism="Homo sapiens"
/mol_type="unassigned DNA"
229-2877
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD22980.1"
/db_xref="GI:18619526"
/translation="MAAAETEQLGVIEFETADCBENIESQDRPKLEPFYVERYSWSQ
LKKLADTRKXVHGMKAPLDFPDKNDPDGPHSDRIYVILAMSGENRNTLFTISEI
PKTINRAAIVLHVMKPLDLFOALDQYMYSEBELAERXIGTIGASIDYHQSG

TFLEAGSGIYHVKGQPGQFTQOPLRPNLVETSPNIRMDPKLPADPDWIAFIHNS
DIWISNVTREERRLTVYHNELANMEEDARSAGVATFVLOEFDFRYSYWCNPKAEIT
PSGKILRLIYEENDESEVEI IHVTPSMLEPRADSFRYPKGTANRVPKPMSEIM
DAEGRIIDVIDKELIQFELLFEGVEYIARAGWTEGKYAWSILLDRSQRQLQIVLIS
PELPIVEDDMERQRLIESVDSVTPPLIYIETTDIWINIHDIHFHVPQSHHEEIEF
IFASECTGFRHLVYKITSILKESKYKRSGLPAPSDPKPIKEBIAITSGEWEVLGR
HGSNIQVDEVRRLVYPEGTKDSPLEHLYVYVNPGBVTRLDRGYSHSCCISOHCD
PFIISKYONKPNHCVSILYKLSPPEDPTCKTKEFWATILDSAGPLPDYTPPELFSRPS
TTGFTLYGMLYKXHDLOPKKYFTVLFIYGGPQVQVNNRPGVKYFRNLNTLASLGYV
VVUIDNRSCHRGKFEKAGPKYKMGQIEIDDDQVEGLQYLAHSRYDFIDLRVGHWSY
GGYLSMALMQRSDIRFVAIAGAPVTLWTFYDTGYTERYMGHPQNEQYVLGSVAMQ
AEKPPSPNRLILLHLGFLDENVHFAITSILLSFLVRAGKPYDLQIYPOERHSRVPES
GEHYELHLLHVLQENLGRIPAAALKVI"

ORIGIN

Query Match 86.9%; Score 1040; DB 6; Length 3143;
Best Local Similarity 89.1%; Pred. No. 9.7e-267;
Matches 1197; Conservative 0; Mismatches 0; Indels 147; Gaps 1;

QY 1 ATTTTGAAGGCCAACAGACTCCCTTTTAGAGCATCACCTGTACGTAGTACGTAA 60
DB 1796 ATTTTGAAGGCCAACAGACTCCCTTTTAGAGCATCACCTGTACGTAGTACGTAA 1855

QY 61 ATCTGGAGAGGTGACAGGCTGACTGACCGTGGCTACTCACATCTTCTGCTGCATCAGTC 120
DB 1856 ATCTGGAGAGGTGACAGGCTGACTGACCGTGGCTACTCACATCTTCTGCTGCATCAGTC 1915

QY 121 AGCACTGTGACTCTTTTATAAGTAAGTATAGTAAACAGAGAAATCCACACTGTGTGCC 180
DB 1916 AGCACTGTGACTCTTTTATAAGTAAGTATAGTAAACAGAGAAATCCACACTGTGTGCC 1975

QY 181 TTTTACAAGCTATCAAGTCTTGAAGTACGACCACTTGGCAAAACAAAGAAATTTTGGGCCA 240
DB 1976 TTTTACAAGCTATCAAGTCTTGAAGTACGACCACTTGGCAAAACAAAGAAATTTTGGGCCA 2035

QY 241 CATTTTGGGATTCAGCAGGCTCTCTTCTGACTATATCTCTCCAGAAATTTTCTCTTTTG 300
DB 2036 CATTTTGGGATTCAGCAGGCTCTCTTCTGACTATATCTCTCCAGAAATTTTCTCTTTTG 2095

QY 301 AAAGTACTTACGATTTACATTTGATGGGATGCTTACAAAGCCTCATGATCTACAGCCTG 360
DB 2096 AAAGTACTTACGATTTACATTTGATGGGATGCTTACAAAGCCTCATGATCTACAGCCTG 2155

QY 361 GAAAGAAATATCTTACTGTGCTGTTTATATATGGTGGTCTCA----- 403
DB 2156 GAAAGAAATATCTTACTGTGCTGTTTATATATGGTGGTCTCTCAGGTGCACTTGGTGAATA 2215
QY 404 ----- 403

DB 2216 ATCGGTTTAAAGGAGTCAAGTATTTCCGTTTGAATACCTAGCTCTCTAGGTTATGTGG 2275
QY 404 ----- 403

DB 2276 TTGTAGTATAGACAAACAGGGATCTCTGTCACCGAGGCTTAAATTTGAAGCGCCCTTA 2335
QY 404 -----GGGTCAAATAGAAATTGACGATCAGGTGGAAGGACTCCAATATCTAGCTT 453
DB 2336 AATATAAAATGGGTCAATAGAAATTGACGATCAGGTGGAAGGACTCCAATATCTAGCTT 2395

QY 454 CTGATATGATTTTCAITGACTTATGATCGTGTGGGCATCCACGGCTGGTCTTATGAGGAT 513
DB 2396 CTGATATGATTTTCAITGACTTATGATCGTGTGGGCATCCACGGCTGGTCTTATGAGGAT 2455

QY 514 ACTCTCCCTGATGCGATTAATGACAGGTGATGATATCTTACAGGTTGCTATTTGCTGGG 573
DB 2456 ACTCTCCCTGATGCGATTAATGACAGGTGATGATATCTTACAGGTTGCTATTTGCTGGG 2515

QY 574 CCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAAAGTTATATGGTGCACC 633
DB 2516 CCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAAAGTTATATGGTGCACC 2575

QY 634 CTGACCAAGTGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAAGAAAGTTCC 693

DB 2576 CTGACCAAGTGAACAGGGCTATTACTTAGGATCTGTGTGCCATGCAAGCAGAAAAGTTCC 2635
QY 694 CCTCTGACCAAAATCGTTTACTCTCTTACATGGTTTCTGCTGATGAGAAATGTCCATTTTG 753
DB 2636 CCTCTGACCAAAATCGTTTACTGCTCTTACATGGTTTCTGCTGATGAGAAATGTCCATTTTG 2695
QY 754 CACATATACCAGTATATTAATCTAGTGGTCTGAAAGCCATATGATTTTACAGA 813
DB 2696 CACATATACCAGTATATTAATCTAGTGGTCTGAAAGCCATATGATTTTACAGA 2755
QY 814 TCTATCTCTCAGGAGAGACACAGCATAGAGTCTCTGAAATCGGGAGAAATATGAACTGC 873
DB 2756 TCTATCTCTCAGGAGAGACACAGCATAGAGTCTCTGAAATCGGGAGAAATATGAACTGC 2815

QY 874 ATCTTTTGCACTACCTTCAAGAAAACCTTGGATCAGTATTGCTGCTCTAAAAGTGATAT 933
DB 2816 ATCTTTTGCACTACCTTCAAGAAAACCTTGGATCAGTATTGCTGCTCTAAAAGTGATAT 2875

QY 934 AATTTTGACCTGTGTAGAACTCTCTGTATACACTGGCTATTAAACCAATGAGGAGGTT 993
DB 2876 AATTTTGACCTGTGTAGAACTCTCTGTATACACTGGCTATTAAACCAATGAGGAGGTT 2935

QY 994 TAATCAACAGAAAACACAGAAATGATCATCACATTTTGTATACCTGCCATGTAACTATAC 1053
DB 2936 TAATCAACAGAAAACACAGAAATGATCATCACATTTTGTATACCTGCCATGTAACTATAC 2995

QY 1054 TCCTGAAAATAAATGTGGTCCATGCAAGGGTCTACGGTTTGTGGTAGTAACTAATACC 1113
DB 2996 TCCTGAAAATAAATGTGGTCCATGCAAGGGTCTACGGTTTGTGGTAGTAACTAATACC 3055

QY 1114 TTAACCCCACTATGCTCAAAATCAAAATGATACATATTTCTTGAGAGACCAGCAATACCAATA 1173
DB 3056 TTAACCCCACTATGCTCAAAATCAAAATGATACATATTTCTTGAGAGACCAGCAATACCAATA 3115

QY 1174 AGAATTTACTAAAAAATAAAAAA 1197
DB 3116 AGAATTTACTAAAAAATAAAAAA 3139

RESULT 8
AR651452
LOCUS AR651452 3120 bp DNA linear PAT 20-APR-2005
DEFINITION Sequence 2 from patent US 6881564.
ACCESSION AR651452
VERSION AR651452.1 GI:62795938
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 3120)
AUTHORS Abbott,C.A. and Gorrell,M.D.
TITLE Dipeptidyl peptidases
JOURNAL Patent: US 6881564-A 2 19-APR-2005;
The University of Sydney; Sydney;
AUX;

FEATURES
source Location/Qualifiers
1..3120
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 86.5%; Score 1036; DB 6; Length 3120;
Best Local Similarity 89.0%; Pred. No. 1.1e-265;
Matches 1193; Conservative 0; Mismatches 0; Indels 147; Gaps 1;

QY 1 ATTTTGAAGGCCAACAAAGACTCCCTTTTAGAGCATCACCTGTACGTAGTACGTAA 60
DB 1781 ATTTTGAAGGCCAACAAAGACTCCCTTTTAGAGCATCACCTGTACGTAGTACGTAA 1840

QY 61 ATCTCTGAGAGGTGACAAAGGCTGACTGACCGTGGCTACTCAATTTCTTGTGTCATCAGTC 120
DB 1841 ATCTCTGAGAGGTGACAAAGGCTGACTGACCGTGGCTACTCAATTTCTTGTGTCATCAGTC 1900

121 AGCACTGTGACTCTCTTTATAGTAAGTATAGTAAACAGAGAAATCCACATGTGTGTCCC 180
1901 AGCACTGTGACTCTCTTTATAGTAAGTATAGTAAACAGAGAAATCCACATGTGTGTCCC 1960
181 TTTTCAAGCTATCAAGTCTGAGATGACCCCACTTGGCAAAACAAAGGAATTTTGGGCCA 240
1961 TTTTCAAGCTATCAAGTCTGAGATGACCCCACTTGGCAAAACAAAGGAATTTTGGGCCA 2020
241 CCATTTTGGATTCAGCAGGTCCTCTCTGCTGCTATATCTCTCCAGAAATTTTCTCTTTTG 300
2021 CCATTTTGGATTCAGCAGGTCCTCTCTGCTGCTATATCTCTCCAGAAATTTTCTCTTTTG 2080
301 AAAGTACTACTGATTTACATGTATGGGATGCTCTACAAGCCTCATGATCTACAGCCTG 360
2081 AAAGTACTACTGATTTACATGTATGGGATGCTCTACAAGCCTCATGATCTACAGCCTG 2140
361 GAAAGAAATATCTCTACTGTGCTGCTTCATATATGGTGGTCTCA----- 403
2141 GAAAGAAATATCTCTACTGTGCTGCTTCATATATGGTGGTCTCA----- 403
404 ----- 403
2201 ATCGTTTAAAGGATCAAGTATTTCCGGCTTGAATACCTAGCCTCTCTAGGTTATGTGG 2260
404 ----- 403
2261 TTGTAGTATAGAACACAGGGGATCTCTGTACCGAGGGCTTAAATTTGAAGGCGCCTTTA 2320
404 -----GGGTCAAAATAGAAATGACGATCAGGTGGAGAGGACTCCAAATATCTAGCTT 453
2321 AATAATAAATGGGTCAAAATAGAAATGACGATCAGGTGGAGAGGACTCCAAATATCTAGCTT 2380
454 CTCGATATGATTTCAITGACTTAGATCGTGTGGGATCCACGGCTGGTCTCATGGAGGAT 513
2381 CTCGATATGATTTCAITGACTTAGATCGTGTGGGATCCACGGCTGGTCTCATGGAGGAT 2440
514 ACCTCTCCCTGATGGCATTAAATGACAGAGTTCAGATATCTTCAGGGTGTCTATTCTCGGG 573
2441 ACCTCTCCCTGATGGCATTAAATGACAGAGTTCAGATATCTTCAGGGTGTCTATTCTCGGG 2500
574 CCCAGTCACTCTGTGGATCTCTATGATACAGGATACACGGAAGCTTATATGGGTCAAC 633
2501 CCCAGTCACTCTGTGGATCTCTATGATACAGGATACACGGAAGCTTATATGGGTCAAC 2560
634 CTGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGCCATGCAAGCAAAAGTTCC 693
2561 CTGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGCCATGCAAGCAAAAGTTCC 2620
694 CCTCTGAACCAATCGTTTACTGCTCTTACATGGTTTCTCTGATGAGAAATGTCATTTTG 753
2621 CCTCTGAACCAATCGTTTACTGCTCTTACATGGTTTCTCTGATGAGAAATGTCATTTTG 2680
754 CACATACAGTATATTACTGAGTTTCTTGTAGTGGGCTGGAAAGCCATATGATTACAGA 813
2681 CACATACAGTATATTACTGAGTTTCTTGTAGTGGGCTGGAAAGCCATATGATTACAGA 2740
814 TCTATCTCAGGAGACACAGATGAAGTTCTGTAATCGGGAGAACATTTGAACTGC 873
2741 TCTATCTCAGGAGACACAGATGAAGTTCTGTAATCGGGAGAACATTTGAACTGC 2800
874 ATCTTTTGCACTACCTTCAAGAAACCTTTGGATCAGTATTCCTGCTCTAAAAGTGATAT 933
2801 ATCTTTTGCACTACCTTCAAGAAACCTTTGGATCAGTATTCCTGCTCTAAAAGTGATAT 2860
934 AATTTTGACCTGTGTAGAACTCTCTGGTATACACTGGCTATTTTAACCAAAATGAGGAGTT 993
2861 AATTTTGACCTGTGTAGAACTCTCTGGTATACACTGGCTATTTTAACCAAAATGAGGAGTT 2920
994 TAATCAACAGAAACACAGAAATTTGATCATCATATTTTGTATACCTGCCATGTAACATCTAC 1053
2921 TAATCAACAGAAACACAGAAATTTGATCATCATATTTTGTATACCTGCCATGTAACATCTAC 2980
1054 TCCTGAAATAAATGTGGTGCCATGACGGGCTACAGGTTTGTGGTAGTATCTAAATACC 1113

2981 TCCTGAAATAAATGTGGTGCCATGACGGGCTACAGGTTTGTGGTAGTATCTAAATACC 3040
1114 TTAACCCCATCATCTCTCAAAATCAAAATGATACATATTCCTGAGAGACCAGCAATACCATA 1173
3041 TTAACCCCATCATCTCTCAAAATCAAAATGATACATATTCCTGAGAGACCAGCAATACCATA 3100
1174 AGAATTACTAAAAA 1193
3101 AGAATTACTAAAAA 3120
BC040203 4535 bp mRNA linear PRI 07-OCT-2003
IMAGE:4824813, complete cds.
BC040203
BC040203.1 GI:26007916
MGC
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 4535)
Strausberg, R.L., Collins, F.S., Wagner, L.H., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shennan, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Shat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Sapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Schetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S., Loquellano, N.A., Peters, G.J.,
McKernan, R.D., Mullah, S.J., Bosak, S.A., McEwan, P.J.,
Akerman, K.J., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Worley, K.C., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Vallal, D.K., Helt, E., Kettman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shvachko, Y.,
Bouffard, G.C., Blakeley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalek, U., Smal, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 4535)
Strausberg, R.
Direct Submission
Submitted (27-NOV-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgaps-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
Toshuyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Louie, H.,
Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
A.N., Gibbs, R.A.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAK Plate: 71 Row: k Column: 23.

FEATURES

source

1..4535
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:42570 IMAGE:4824813"
/tissue_type="Testis"
/clone_lib="NIH_MGC_97"
/lab_host="DH10B"
/note="Vector: pBluescript"
1..4535
/gene="DPP8"
/note="synonyms: FLJ20283, FLJ14920, MGC26191"
/db_xref="GeneID:54878"
/db_xref="MIM:606819"
1629..4277
/gene="DPP8"
/codon_start=1
/product="Dpp8 protein"
/protein_id="AAH40203.1"
/db_xref="GI:26007917"
/db_xref="GeneID:54878"
/db_xref="MIM:606819"
/translations:"MAAAMETQLGVETADCEENIESODRPLKPFVVERYSWSQ
LKLLADTRYKHGYMMAKAPHDFMFVXRNDPDGPHSDRIYYLAMSGENRENTLFYSEI
PKTINRAAAMLSWSKPLDLFOALDYGMYREBELRKRRIKLGPIADYIADYHQSG
TFLFOAGSGIYHVKGDPGPGFTQPLRPNLVETSPNIRMDPKLCPADPDWIAFIHGN
DIMISNIVREERLYVYNELANMEEDARSAGVATFVLOEFPDRISGYWPCAEIT
PSGKILRLIDYDESEVEIIHVTSFMLETRADSPRIKPTGTANPKVTKWSEIMI
DAEGRIIDYDEKLIQFFELFEGVEYIARAGWTPEGKYAWSILLDRSQRLOIVLIS
PELFIPEDDVMORORLIESVPSVTPLIIVYEETDIIWINIHDIHFVFPQSHHEIEF
IFASECTGPHLYKITSILKSKYKRSGLPAPSPDFKPIKEEIAITSGEWEVLGR
HGSNIQVDEVRILYFECKDSPLEHLVYVSYNPNCEVETLDRGSHSCCISQHCD
FTFISYQNKPNHCVSLYKLSSPEDDPTCKEFWATILDSAGLPDYITPEIFSES
TTGFTLYMLKPHDLQPGKYPTLFIYGGPQVLNNRPFKYFRNLNTLASLGV
VVIDNRGSHRGLKEGAFKMGQIEIDQVEGLQYLASRYDFIDLRVGHWSY
GGVLSLWALMSRDI PRVAIAGAPVTLMFYDTGVTYRYMGHPDNQBGYYLGSVAMQ
AEKFPSEPNRLLHLHGFLENVHFAHTSILLSFLVRAGKPYDLQIYQERHSIRVPES
GEHYELHLHYLOENLGSRIAALKVI"
3207..4196
/gene="DPP8"
/note="DAP2; Region: Dipeptidyl
aminopeptidases/acylaminoacyl-peptidases [Amino acid
transport and metabolism]"
/db_xref="CDD:COG1506"

misc_feature

ORIGIN

Query Match 86.5%; Score 1036; DB 8; Length 4535;
Best Local Similarity 89.0%; Pred. No. 1.1e-265;
Matches 1193; Conservative 0; Mismatches 0; Indels 147; Gaps 1;
QY 1 ATTTTGAGGCACCAAGACTCCCTTTAGAGCATCACCTGACCTAGCTAGCTACGTAA 60
DB 3196 ATTTTGAGGCACCAAGACTCCCTTTAGAGCATCACCTGACCTAGCTAGCTAA 3255
QY 61 ATCTCTGAGAGGTGACAAAGCTGACTGACCGTGGCTACTCACATCTTGTGTCATCAGTC 120
DB 3256 ATCTCTGAGAGGTGACAAAGCTGACTGACCGTGGCTACTCACATCTTGTGTCATCAGTC 3315
QY 121 AdCACTGTGACTCTTTTATAAGTAGATAGTAAACAGAGAATCCACACTGTGTGTCCC 180
DB 3316 AdCACTGTGACTCTTTTATAAGTAGATAGTAAACAGAGAATCCACACTGTGTGTCCC 3375
QY 181 TTtACAAGCTATCAAGTCTGAAAGATGACCCAACTTGCAAAACAAAGGAATTTTGGGCCA 240
DB 3376 TTtACAAGCTATCAAGTCTGAAAGATGACCCAACTTGCAAAACAAAGGAATTTTGGGCCA 3435
QY 241 CCAATTTGGATTAGCAGAGTCTCTTCTCGACTATACCTCCAGAAATTTCTCTTTTG 300
DB 3436 CCAATTTGGATTAGCAGAGTCTCTTCTCGACTATACCTCCAGAAATTTCTCTTTTG 3495
QY 301 AAAGTACTACTGGATTTACATTTGATGGGATGCTCTACAAGCCTCATGATCTACAGCCTG 360

RESULT 10

AR631284
LOCUS AR631284 4309 bp DNA linear PAT 14-FEB-2005
DEFINITION Sequence 14 from patent US 6844180.
ACCESSION AR631284
KEYWORDS AR631284.1 GI:59770929
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 4309)
AUTHORS Qi,S., Akineanya,K.O., Riviere,P.J.M. and Junien,J.-L.
TITLE Serine protease genes related to DPPIV
JOURNAL Patent: US 6844180-A 14 18-JAN-2005;
Ferring BV;;
NLX;
FEATURES
source Location/Qualifiers
1..4309
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 86.3%; Score 1033.2; DB 6; Length 4309;
Best Local Similarity 88.7%; Pred. No. 6.2e-265;
Matches 1194; Conservative 0; Mismatches 3; Indels 149; Gaps 1;
1
1 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 60
2 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 120
3 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 180
4 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 240
5 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 300
6 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 360
7 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 420
8 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 480
9 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 540
10 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 600
11 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 660
12 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 720
13 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 780
14 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 840
15 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 900
16 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 960
17 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 1020
18 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 1080
19 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 1140
20 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 1200
21 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 1260
22 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 1320
23 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 1380
24 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 1440
25 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 1500
26 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 1560
27 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 1620
28 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 1680
29 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 1740
30 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 1800
31 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 1860
32 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 1920
33 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 1980
34 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 2040
35 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 2100
36 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 2160
37 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 2220
38 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 2280
39 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 2340
40 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 2400
41 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 2460
42 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 2520
43 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 2580
44 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 2640
45 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 2700
46 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 2760
47 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 2820
48 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 2880
49 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 2940
50 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 3000
51 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 3060
52 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 3120
53 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 3180
54 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 3240
55 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 3300
56 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 3360
57 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 3420
58 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 3480
59 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 3540
60 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 3600
61 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 3660
62 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 3720
63 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 3780
64 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 3840
65 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 3900
66 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 3960
67 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 4020
68 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 4080
69 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 4140
70 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 4200
71 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 4260
72 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 4320
73 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 4380
74 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 4440
75 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 4500
76 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 4560
77 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 4620
78 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 4680
79 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 4740
80 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 4800
81 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 4860
82 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 4920
83 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 4980
84 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 5040
85 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 5100
86 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 5160
87 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 5220
88 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 5280
89 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 5340
90 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 5400
91 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 5460
92 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 5520
93 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 5580
94 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 5640
95 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 5700
96 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 5760
97 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 5820
98 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 5880
99 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 5940
100 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCAGCTGTAGCTAGTTCAGTTACGTAA 6000

572 GGGCCAGTCACTCTGTGGATCTTCTATATACAGATACACGGAACGTTATATGGGTCA 631
1981 GGGCCAGTCACTCTGTGGATCTTCTATATACAGATACACGGAACGTTATATGGGTCA 2040
632 CCCTGACAGAAATGAACAGGGCTATTACTTAGATCTGTGGCCATGCAAGCAGAAAAAGTT 691
2041 CCCTGACAGAAATGAACAGGGCTATTACTTAGATCTGTGGCCATGCAAGCAGAAAAAGTT 2100
692 CCCCTCTGAACCAAAATCGTTTACTGCTCTTACATGTTTCTCTGGATGAGATGTCATTT 751
2101 CCCCTCTGAACCAAAATCGTTTACTGCTCTTACATGTTTCTCTGGATGAGATGTCATTT 2160
752 TGCATACACAGATATATTACTGAGTTTCTTGTAGGGCTGGAAGCCATATGATTTACA 811
2161 TGCATACACAGATATATTACTGAGTTTCTTGTAGGGCTGGAAGCCATATGATTTACA 2220
812 GATCTATCTCTCAGGAGACACAGATTAAGAGTTCTGAATCGGAGAACATTTATGAAT 871
2221 GATCTATCTCTCAGGAGACACAGATTAAGAGTTCTGAATCGGAGAACATTTATGAAT 2280
872 GCATCTTTTGCATCTACCTTCAAGAAAACTTGGATCAGTATTGCTGCTCTAAAGTGAT 931
2281 GCATCTTTTGCATCTACCTTCAAGAAAACTTGGATCAGTATTGCTGCTCTAAAGTGAT 2340
932 ATATTTTGGACCTGTGTAGAACTCTCTGTATACACTGGCTATTTAACCAATGAGGAG 991
2341 ATATTTTGGACCTGTGTAGAACTCTCTGTATACACTGGCTATTTAACCAATGAGGAG 2400
992 TTTTAATCAACAGAAAAACAGAAATGATCATCATCATTTTGTACTGCGCATGTAACATCT 1051
2401 TTTTAATCAACAGAAAAACAGAAATGATCATCATCATTTTGTACTGCGCATGTAACATCT 2460
1052 ACTCTGAAAATTAATGTGTGGTGCATGAGGGGTCTACGGTTTGTGTAGTAATCTAATA 1111
2461 ACTCTGAAAATTAATGTGTGGTGCATGAGGGGTCTACGGTTTGTGTAGTAATCTAATA 2520
1112 CTTTAAACCCCAATGCTCAAAATCAAAATGATACATATTTCTGAGAGACCCAGCAATACCA 1171
2521 CTTTAAACCCCAATGCTCAAAATCAAAATGATACATATTTCTGAGAGACCCAGCAATACCA 2580
1172 TAAGAAATTAATAAAAAAAAAAAAAA 1197
2581 TAAGAAATTAATAAAAAAAAAAAAAA 2606
RESULT 11
AX608737
LOCUS AX608737 4309 bp DNA linear PAT 17-FEB-2003
DEFINITION Sequence 14 from Patent WO0231134.
ACCESSION AX608737
VERSION AX608737.1 GI:28404304
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE 1
AUTHORS Qi,S., Akineanya,K.O., Riviere,P.J. and Junien,J.L.
TITLE Novel serine protease genes related to dppiv
JOURNAL Patent: WO 0231134-A 14 18-APR-2002;
Ferring BV (NL)
FEATURES
source Location/Qualifiers
1..4309
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 86.3%; Score 1033.2; DB 6; Length 4309;
Best Local Similarity 88.7%; Pred. No. 6.2e-265;
Matches 1194; Conservative 0; Mismatches 3; Indels 149; Gaps 1;

QY 1 ATTTTGAAGGCAACAAAGACTCCCTTTTAGAGATCACTGTAGCTAGTCACTGTA 60
DB 1261 ATTTTGAAGGCAACAAAGACTCCCTTTTAGAGATCACTGTAGCTAGTCACTGTA 1320
QY 61 ATCTGTGAGAGGTGACAAAGGCTGACTGACCGTGGTACTCACTTTCTGTGTCATCAGTC 120
DB 1321 ATCTGTGAGAGGTGACAAAGGCTGACTGACCGTGGTACTCACTTTCTGTGTCATCAGTC 1380
QY 121 AGCACTGTGATCTCTTTTATAGTAAGTATAGTAACCAAGAAATTCCTGTTGTC 180
DB 1381 AGCACTGTGATCTCTTTTATAGTAAGTATAGTAACCAAGAAATTCCTGTTGTC 1440
QY 181 TTTTACAGCTATCAAGTCTGAGATGACCAACTTGCAGAAAGAAATTTTGGGCCA 240
DB 1441 TTTTACAGCTATCAAGTCTGAGATGACCAACTTGCAGAAAGAAATTTTGGGCCA 1500
QY 241 CCATTTTGGATTCAGCAGGTCCTCTTCTGACTATATCTCTCCAGAAATTTCTCTTTTG 300
DB 1501 CCATTTTGGATTCAGCAGGTCCTCTTCTGACTATATCTCTCCAGAAATTTCTCTTTTG 1560
QY 301 AAAGTACTACTGGATTTACATTTGATGGATGCTCTACAAGCCTCATGATCTACAGCCTG 360
DB 1561 AAAGTACTACTGGATTTACATTTGATGGATGCTCTACAAGCCTCATGATCTACAGCCTG 1620
QY 361 GAAAGAAATATCCTACTGTGCTGTTTCATATATGTTGGTCTC----- 402
DB 1621 GAAAGAAATATCCTACTGTGCTGTTTCATATATGTTGGTCTCCTCAGGTGCGAGTTGGTGA 1680
QY 403 ----- 402
DB 1681 TAATCGGTTTAAAGGAGTCAAGTATTTCCGTTTGAATACCCCTAGCCTCTCTAGGTTATGT 1740
QY 403 ----- 402
DB 1741 GGTTGTAGTAGACAAACAGGGGATCCTGTCAACGAGGGCTTAAATTTGAAGCGGCTT 1800
QY 403 -----AGGGTCAAAATAGAAATTGACGATCAGGTGGAAGGACTCCCAATATCTAGC 451
DB 1801 TAAATATAAATGGGTCAAAATAGAAATTGACGATCAGGTGGAAGGACTCCCAATATCTAGC 1860
QY 452 TTCTCGATATGATTTTCAATGACTTAGATCGTGTGGGCATCCACGGCTGGTCTCTATGAGG 511
DB 1861 TTCTCGATATGATTTTCAATGACTTAGATCGTGTGGGCATCCACGGCTGGTCTCTATGAGG 1920
QY 512 ATACCTCTCCTGTATGSCATTAATGCAGAGGTGAGATATCTTCAGGGTGTCTATTTGCTGG 571
DB 1921 ATACCTCTCCTGTATGSCATTAATGCAGAGGTGAGATATCTTCAGGGTGTCTATTTGCTGG 1980
QY 572 GGGCCCAAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAACTTATATGAGGTCA 631
DB 1981 GGGCCCAAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAACTTATATGAGGTCA 2040
QY 632 CCTGTACAGATGAAACAGGGCTATTACTTAGATCTGTGGCCATGCAAGCAGAAAGTT 691
DB 2041 CCTGTACAGATGAAACAGGGCTATTACTTAGATCTGTGGCCATGCAAGCAGAAAGTT 2100
QY 692 CCCCTGTGAACCAATCGTTTACTGTCTTACATGTTTCTCGATGAGATGTCCTATTT 751
DB 2101 CCCCTGTGAACCAATCGTTTACTGTCTTACATGTTTCTCGATGAGATGTCCTATTT 2160
QY 752 TdCATATACAGTATATTTACTGAGTGTCTTTTGTAGGGGCTGGAAGCCATATGATTACA 811
DB 2161 TdCATATACAGTATATTTACTGAGTGTCTTTTGTAGGGGCTGGAAGCCATATGATTACA 2220
QY 812 GATCTATCTCCTCAGGAGAGACACAGCATAAGAGTTCTGTAATCGGAGAACATTATGAAT 871
DB 2221 GATCTATCTCCTCAGGAGAGACACAGCATAAGAGTTCTGTAATCGGAGAACATTATGAAT 2280
QY 872 GCATCTTTTGCATCTCTCAAGAAACCTTTGATACAGTATGCTGTCTTAAAGTGAT 931
DB 2281 GCATCTTTTGCATCTCTCAAGAAACCTTTGATACAGTATGCTGTCTTAAAGTGAT 2340
QY 932 ATAAATTTTACCTGTGTAGAACTCTCTGTGTATACACTGGCTATTTAAACCAATAGGAGG 991

DB 2341 ATAATTTTACCTGTGTAGAACTCTCTGGTATACACTGGCTATTTAACCAATAGGAGG 2400
QY 992 TTTTAAATCAACAGAAACACAGAAATTTGATCATCACATTTTGTACCTGCCATGTAACTCT 1051
DB 2401 TTTTAAATCAACAGAAACACAGAAATTTGATCATCACATTTTGTACCTGCCATGTAACTCT 2460
QY 1052 ACTCCTGAAATTAATCTGTGTCATGACGGGTCTACGGTTTGTGGTAGTAATCTAATA 1111
DB 2461 ACTCCTGAAATTAATCTGTGTCATGACGGGTCTACGGTTTGTGGTAGTAATCTAATA 2520
QY 1112 CTTTAAACCCCATGCTGCTCAAAATCAATGATACATATTTCTCGAGAGACCCAGCAATACCA 1171
DB 2521 CTTTAAACCCCATGCTGCTCAAAATCAATGATACATATTTCTCGAGAGACCCAGCAATACCA 2580
QY 1172 TAAGATTTACTTAAATAAAAAAAAA 1197
DB 2581 TAAGATTTACTTAAATAAAAAAAAA 2606
RESULT 12
AR631283
LOCUS AR631283 4829 bp DNA linear PAT 14-FEB-2005
DEFINITION Sequence 12 from patent US 6844180.
ACCESSION AR631283
VERSION AR631283.1 GI:59770927
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 4829)
AUTHORS Oi, S., Akinsanya, K.O., Riviere, P.J.M. and Junien, J.-L.
TITLE Serine protease genes related to DPPIV
JOURNAL Patent: US 6844180-A 12 18-JAN-2005;
Ferring BV;
NLX;
FEATURES
source Location/Qualifiers
1..4829
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 86.3%; Score 1033.2; DB 6; Length 4829;
Best Local Similarity 88.7%; Pred. No: 6.1e-265;
Matches 1194; Conservative 0; Mismatches 3; Indels 149; Gaps 1;
QY 1 ATTTTGAAGGCAACAAAGACTCCCTTTTAGAGATCACTGTAGCTAGTCACTGTA 60
DB 1781 ATTTTGAAGGCAACAAAGACTCCCTTTTAGAGATCACTGTAGCTAGTCACTGTA 1840
QY 61 ATCTGTGAGAGGTGACAAAGGCTGACTGACCGTGGTACTCACTTTCTGTGTCATCAGTC 120
DB 1841 ATCTGTGAGAGGTGACAAAGGCTGACTGACCGTGGTACTCACTTTCTGTGTCATCAGTC 1900
QY 121 AGCACTGTGATCTCTTTTATAGTAAGTATAGTAACCAAGAAATTCCTGTTGTC 180
DB 1901 AGCACTGTGATCTCTTTTATAGTAAGTATAGTAACCAAGAAATTCCTGTTGTC 1960
QY 181 TTTTACAGCTATCAAGTCTGAGATGACCCAACTTGCAGAAAGAAATTTTGGGCCA 240
DB 1961 TTTTACAGCTATCAAGTCTGAGATGACCCAACTTGCAGAAAGAAATTTTGGGCCA 2020
QY 241 CCATTTTGGATTCAGCAGGTCCTCTTCTGACTATATCTCTCCAGAAATTTTCTCTTTTG 300
DB 2021 CCATTTTGGATTCAGCAGGTCCTCTTCTGACTATATCTCTCCAGAAATTTTCTCTTTTG 2080
QY 301 AAAGTACTACTGGATTTTACATTTGATGGATGCTCTACAAGCCTCATGATCTACAGCCTG 360
DB 2081 AAAGTACTACTGGATTTTACATTTGATGGATGCTCTACAAGCCTCATGATCTACAGCCTG 2140
QY 361 GAAAGAAATATCCTACTGTGCTGTTTCATATATGTTGGTGGTCTCT----- 402
DB 2141 GAAAGAAATATCCTACTGTGCTGTTTCATATATGTTGGTGGTCTCTCTCAGGTGCGAGTTGGTGA 2200

Db 2850 AATTTGACCTGTAGAACTCTCTGGTATACACTGGCTATTATTAACCAATGAGAGGTT 2909

Qy 994 TAATCACAGAAACACAGAAATTGATCATCATCATTTTGTATCTCCCATGTAAACATCTAC 1053

Db 2910 TAATCACAGAAACACAGAAATTGATCATCATCATTTTGTATCTCCCATGTAAACATCTAC 2969

Qy 1054 TCCTGAAATAAATGTGGTCCCATGAGGGGCTACAGGTTTGTGTAGTAACTAAATACC 1113

Db 2970 TCCTGAAATAAATGTGGTCCCATGAGGGGCTACAGGTTTGTGTAGTAACTAAATACC 3029

Qy 1114 TTAACCCCATGCTCAAAATCAAAATGATACATATTTCTCTGAGAGACCCAGCAATACCATA 1173

Db 3030 TTAACCCCATGCTCAAAATCAAAATGATACATATTTCTCTGAGAGACCCAGCAATACCATA 3089

Qy 1174 AGAATTACTTAAAAAAA 1190

Db 3090 AGAATTACTTAAAAAAA 3106

RESULT 15

BC030688 3125 bp mRNA linear PRI 28-JUL-2005

LOCUS Homo sapiens dipeptidylpeptidase 8, transcript variant 4, mRNA

DEFINITION (cDNA clone MGC:26191 IMAGE:4822550), complete cds.

ACCESSION BC030688

VERSION BC030688.2 GI:34190028

KEYWORDS MGC.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.

AUTHORS 1 (bases 1 to 3125)

1 Strausberg, R.L., Feingold, E.A., Grouse, L.H., Berge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heieh, F., Datchenko, L., Marudina, K., Farmer, A.A., Rubin, G.M., Hong, L., Scapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Cesaevant, T.L., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullah, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Hale, S.J., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Touchman, J.W., Green, E.D., Bouffard, G.G., Blakesley, R.W., Young, A.C., Shevchenko, Y., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smailus, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.

CONSR TM Mammalian Gene Collection Program Team

TITLE Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

REFERENCE Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

PUBMED 12477932

AUTHORS 2 (bases 1 to 3125)

NIH MGC Project

Direct Submission

JOURNAL Submitted (24-MAY-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Bethesda, MD 20892-2590, USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>

COMMENT On Aug 25, 2003 this sequence version replaced gi:21265132.

Contact: MGC help desk

Email: cgabbe-x@mail.nih.gov

Tissue Procurement: Miklos Palkovits, M.D., Ph.D.

cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki Toshiyuki and Piero Carninci (RIKEN)

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Institute for Systems Biology

<http://www.systemsbio.org>

contact: amadansystemsbiology.org

Anup Madan, Jessica Fahey, Brin Helton, Mark Kettenan, Anuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAC Plate: 33 Row: d Column: 5

This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 37577092.

Location/Qualifiers

1. 3125

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="MGC:26191 IMAGE:4822550"

/tissue_type="Testis"

/clone_lib="NIH MGC 97"

/lab_host="DH10E"

/note="Vector: pBluescriptR"

1. 3125

/gene="DPP8"

/notes="synonyms: FLJ20283, FLJ14920, MGC26191, MSTP141"

/db_xref="GeneID:54878"

/db_xref="MIM:606819"

319..2862

/gene="DPP8"

/codon_start=1

/product="dipeptidyl peptidase 8, isoform 4"

/protein_id="AAH30688.3"

/db_xref="GI:54114505"

/db_xref="GeneID:54878"

/db_xref="MIM:606819"

/translations="MWKSEQWKIKSGKCNMAANETQLGVETFDACBENTESQD RPKLEFFYVERYSQKKLADATRYKHGMKAPHDFMFVKNRDPDGHSDRIYIL AMGSENNRFLYSEIPKTNRAALMISWKLDFLDFORTLDYDNGYSEELLERKR IGTVASDYHQSGTFLFQSGIYHVKGQPGQFTQQLPRLNVLVETSCPNRMOP KLCPADWIAFTHSNDIMVSNIVREERRLYVHNEANMEEDARSAGVATFVLQEE FDRYSGWCPKAEATTPSGGKILRLYIENDESEVEIIVHTSPMLERADSPRYPKT GTANPKVTFKSEIMIDAGRIIDVDELKOPFEILFEGVEYIARAGTDEPGYAWG ILLRSQTRQLVILSELPFIPEVDVMEQRILISLKSQSVKSSGGLVPAESDFKPI IDIHFVPSQHEEIEFISECKTGRHLIKITSLKESKYSRSGGLVPAESDFKPI KESIAITSEWELVGHNSIQVDEVRLLVPEGKDSPLFELHLYVVSYNVPDFKRL TDTYSHSCISQHCDFISYKSNQKPHCVSLYKLSPEDDPTCKKEFWATILDA GFLPDTYTPFIETFTGTTLTGMLYKPHDLQPGKPYTLVLYFGQVQLVNNRFK GVYFRLNTLASLGYVVVIDNRGSHRGLKPEGAFKYMVAIAGAPVTLWIFDVTGY TERYMGHPDQNEQGYLVGSVAMQAEKFPSEPNLLHLLHFLDENVHFAHTSILLSFLV RAGKPYDLQIYPOERHSIRVPESGSHYELHLLHLYQENLGSRIAALKVI"

ORIGIN

Query Match 84.9%; Score 1016.2; DB 8; Length 3125;

Best Local Similarity 91.7%; Pred. No. 2.2e-260;

Matches 1098; Conservative 0; Mismatches 93; Indels 6; Gaps 2;

Qy 1 ATTTTGAAGCCACCAAGACTCCCTTTAGAGCATCACCTGTACGTACGTACGTACGTA 60

Db 1934 ATTTTGAAGCCACCAAGACTCCCTTTAGAGCATCACCTGTACGTACGTACGTA 1993

Qy 61 ATCTGGAGAGGTGACAGAGCTGACGACCGTGTACTCACATCTCTTGTGTCATCAGTC 120

Db 1994 ATCTGGAGAGGTGACAGAGCTGACGACCGTGTACTCACATCTCTTGTGTCATCAGTC 2053

Qy 121 AGCACTGTGACTCTTTTATAGTAAGTATAGTAACAGAGAAATCCACATCTGTGTGTC 180

Db 2054 AGCACTGTGACTCTTTTATAGTAAGTATAGTAACAGAGAAATCCACATCTGTGTGTC 2113

Qy 181 TTTTAAAGCTATCAAGTCTCGAAGATGACCCAACTTGGCAAAACAAAGAAATTTGGGCA 240

Db 2114 TTTTAAAGCTATCAAGTCTCGAAGATGACCCAACTTGGCAAAACAAAGAAATTTGGGCA 2173

Qy 241 CCATTTTGGATTACAGAGGTCTCTTCTCTGACTATATCTCTCCAGAAATTTCTCTTTTG 300

Db 2174 CCATTTTGGATTACAGAGGTCTCTTCTCTGACTATATCTCTCCAGAAATTTCTCTTTTG 2233

Qy 301 AAGTACTACTGGATTATACATTTGATGGGATGCTCTTACAAGCCTCATGATCTACAGCCTG 360

Db 2234 AAAGTACTACTGGATTTACATTTGATGGGATGCTCTCAAGCCTCATGATCTACAGCCTG 2293
Qy 361 GAAAGAAATATCTACTGTCTGTTCATATATGTTGGTCCCTCAGGGTCAATAGAAATTG 420
Db 2294 GAAAGAAATATCTACTGTCTGTTCATATATGTTGGTCCCTCAGGGTCAAGTTGG--TGA 2350
Qy 421 ACGATCAGGTGGAGGACTCCCAATATCTAGCTTCTCGATATGATTTTCATTGACTTAGATC 480
Db 2351 ATATCGGTTTAAAGAGGTCAAGTATTTCCGCT---TGAATACCCCTAGCCTCTCTAGGTT 2407
Qy 481 GTGTGGGCATCCACGGCTGGTCTCTATGAGGAGTACTCTCTCGATGGCATTAATGCAGA 540
Db 2408 ATGTGGTTGTAGTATAGACACAGGGGATCTGTACGGAGGGCTTAATTTGAAGGCG 2467
Qy 541 GGTCAAGATATCTTTCAAGGTTGTCTATGTGGGGCCCCAGTCACTCTGTGGATCTTCTATG 600
Db 2468 CCTTTAAATATAAATGGTTGCTATTTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATG 2527
Qy 601 ATACAGGATACACGGAACGTTATATATGTTGGTCACTGTACAGAAATGAACAGGGCTATTACT 660
Db 2528 ATACAGGATACACGGAACGTTATATGTTGGTCACTGTACAGAAATGAACAGGGCTATTACT 2587
Qy 661 TAGGATCTGTGGCCATGCAAGCAGAGAAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCT 720
Db 2588 TAGGATCTGTGGCCATGCAAGCAGAGAAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCT 2647
Qy 721 TACATGGTTTCTGGATGAGATGCCATTTTGGACATACCAGTATATTACTGAGTTTCT 780
Db 2648 TACATGGTTTCTGGATGAGATGCCATTTTGGACATACCAGTATATTACTGAGTTTCT 2707
Qy 781 TAGTGGGGCTGGAAGGCCATATGATTTACAGATCTATCTCAGGAGAGACACAGCATAA 840
Db 2708 TAGTGGGGCTGGAAGGCCATATGATTTACAGATCTATCTCAGGAGAGACACAGCATAA 2767
Qy 841 GAGTTCCTGAATCGGAGAACATTTATGAACTGCTCTTTTGGCACTTACCTTCAAGAAAACC 900
Db 2768 GAGTTCCTGAATCGGAGAACATTTATGAACTGCTCTTTTGGCACTTACCTTCAAGAAAACC 2827
Qy 901 TTGGATCACGTATTTGTCTCTTAAAGTGATATAATTTTGGACCTGTGTAGAACTCTCTGG 960
Db 2828 TTGGATCACGTATTTGTCTCTTAAAGTGATATAATTTTGGACCTGTGTAGAACTCTCTGG 2887
Qy 961 TATACACTGGCTATTTAAACCAATGAGGAGGTTTAAATCAACAGAAAACACAGAAATTGATC 1020
Db 2888 TATACACTGGCTATTTAAACCAATGAGGAGGTTTAAATCAACAGAAAACACAGAAATTGATC 2947
Qy 1021 ATACATTTTGTATACCTGCCATGTACATCTACTCTGAAATAAATGGTGGCCATGCA 1080
Db 2948 ATACATTTTGTATACCTGCCATGTACATCTACTCTGAAATAAATGGTGGCCATGCA 3007
Qy 1081 GGGGTCTACGGTTTGTGGTAGTAACTTAATACCTTAACCCACATGCTCAAAATCAAAATG 1140
Db 3008 GGGGTCTACGGTTTGTGGTAGTAACTTAATACCTTAACCCACATGCTCAAAATCAAAATG 3067
Qy 1141 ATACATTTTCTCAGAGACCCAGCAATACCAATAGAAATTACTTAAAAAAA 1197
Db 3068 ATACATTTTCTCAGAGACCCAGCAATACCAATAGAAATTACTTAAAAAAA 3124

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.
OM nucleic - nucleic search, using sw model
Run on: May 4, 2006, 02:50:44 ; Search time 674.127 Seconds
(without alignments)
11834.025 Million cell updates/sec
Title: US-10-825-632-4
Perfect score: 1197
Sequence: 1 atttgaaggccacaaagac.....ttactaataaaaaaaaaa 1197
Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0
Searched: 4996997 seqs, 3332346308 residues
Total number of hits satisfying chosen parameters: 9993994
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :			
N_Geneseq_21.*			
1:	Geneseqn1980s.*		
2:	Geneseqn1990s.*		
3:	Geneseqn2000s.*		
4:	Geneseqn2001as.*		
5:	Geneseqn2001bs.*		
6:	Geneseqn2002as.*		
7:	Geneseqn2002bs.*		
8:	Geneseqn2003as.*		
9:	Geneseqn2003bs.*		
10:	Geneseqn2003cs.*		
11:	Geneseqn2003ds.*		
12:	Geneseqn2004as.*		
13:	Geneseqn2004bs.*		
14:	Geneseqn2005s.*		

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match	Length DB ID Description
1	1197	100.0	1197 4 AAC85695 Nucleotid
2	1179	98.5	4676 6 ABK83331
3	1040	86.9	2083 7 ADRA1222 Human CD-
4	1040	86.9	2797 12 ADL13374 Human ste
5	1040	86.9	3143 6 AAH99934
6	1036	86.5	3120 4 AAC85694
7	1036	86.5	3120 6 AAD38956
8	1033.2	86.3	4309 6 ABK83328
9	1033.2	86.3	4829 6 ABK83327
10	1033	86.3	3106 6 ABK12892
11	1030	86.0	2929 10 ACA92425
12	1021	85.3	1821 6 ABV76411
13	1013.6	84.7	4685 6 ABK83332
14	1005.6	84.0	2952 10 ACA92421
15	999.6	83.5	2161 4 AAH15009
16	897.6	75.0	2251 10 ADE79035
17	881	73.6	1669 4 AAC85696
18	879	73.4	2668 6 ABN59775
19	879	73.4	2842 6 ABN59774

20	873	72.9	4523	6	ABK83325
21	863.6	72.1	2463	10	ACA92424
22	863.6	72.1	2510	6	AAD23843
23	840.8	70.2	2702	13	ADT04072
24	791.4	66.1	873	4	AAF81719
25	791.4	66.1	925	6	ABL90148
26	778	65.0	2649	14	ADV43981
27	776.4	64.9	2649	8	ABX12255
28	776.4	64.9	2671	6	ABK83322
29	775	64.7	2643	6	AAH99935
30	633.4	52.9	823	6	ABK30401
31	619	51.7	2349	14	ADV43982
32	523.6	43.7	587	4	AAH12830
33	508	42.4	2696	12	ADI16386
34	459	38.3	1083	4	AAC85697
35	380.4	31.8	561	4	AAJ00876
36	326.8	27.3	612	6	ABN86324
37	308.8	25.8	502	9	ACH32373
38	237.2	19.8	3287	6	AAD38955
39	214.4	17.9	2660	6	ADI16689
40	214.4	17.9	2660	6	ADI16687
41	212.8	17.8	662	4	AAK92083
42	212.8	17.8	662	4	AAK93366
43	212.8	17.8	662	12	ADI29793
44	212.8	17.8	662	12	ADI28510
45	212.8	17.8	1048	4	AAS41004

ALIGNMENTS

RESULT 1
AAC85695
ID AAC85695 standard; cDNA; 1197 BP.

XX AAC85695;

XX 29-JUN-2001 (first entry)

DE Nucleotide sequence encoding human DPP8 524Phe-657Gln-707Gly-882Ile.

XX Human; dipeptidyl aminopeptidase; DPP8; prollyl oligopeptidase;
KW dipeptidyl peptidase; DPPIV; T cell; cleavage; diarrhoea;
KW growth hormone deficiency; glucose level; mucosal regeneration;
KW non-insulin dependent diabetes mellitus; glucose intolerance;
KW immunosuppression; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

FT CDS 3..935

FT /*tag= a /product= "Human DPP8 524Phe-657Gln-707Gly-882Ile"

XX WO200119866-A1.

PN 22-MAR-2001.

XX 11-SEP-2000; 2000WO-AU001085.

XX 10-SEP-1999; 99AU-00002762.

PR 18-FEB-2000; 2000AU-00005709.

XX (UNSY) UNIV SYDNEY.

XX PI Abbott CA, Gorell MD;

XX WPI; 2001-281520/29.

XX P-PSDB; AAB47188.

XX New human dipeptidyl aminopeptidase (DPP8) useful for cleaving
PT substrates, identifying inhibitors of DPP8 catalytic activity which have
PT therapeutic uses, and for detecting activated T cells.

XX The present invention relates to the isolation of novel human serine
CC proteases referred to as dipeptidyl peptidase IV (DPPIV)-related proteins
CC (DPRP). The dipeptidyl peptidase IV-related proteins (DPRP) and nucleic
CC acids encoding them are useful for treating infections such as fungal,
CC bacterial, protozoan and viral infections, particularly infections caused
CC by human immunodeficiency virus (HIV-1 or HIV-2), pain, diabetes,
CC precocious puberty, infertility, obesity, anorexia, bulimia, Parkinson's
CC disease, acute heart failure, hypotension, hypertension, urinary
CC retention, osteoporosis, angina pectoris, stroke, ulcers, asthma,
CC allergies, cancers, migraine, vomiting, psychotic and neurological
CC disorders (e.g. anxiety, dementia, or schizophrenia), and dyskinesias.
CC These may also be used in discovering therapeutic agents for the
CC treatment of reproductive, inflammatory and metabolic disorders. ABK83322
CC -ABK83343 encode human DPRP proteins
XX
SQ Sequence 4676 BP; 1424 A; 859 C; 979 G; 1414 T; 0 U; 0 Other;

Query Match 98.5%; Score 1179; DB 6; Length 4676;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1193; Conservative 0; Mismatches 0; Indels 4; Gaps 1;

QY 1 ATTTGAAGGCAACCAAGACTCCCTTTAGAGCATCACCTGTAGTACGTTACGTAA 60
DB 1781 ATTTGAAGGCAACCAAGACTCCCTTTAGAGCATCACCTGTAGTACGTTACGTAA 1840

QY 61 ATCTGTGAGAGTGACAGGCTGACTGACCGTGGTACTCACATTTCTGTCATCAGTC 120
DB 1841 ATCTGTGAGAGTGACAGGCTGACTGACCGTGGTACTCACATTTCTGTCATCAGTC 1900

QY 121 AGCACTGTGACTTCTTTAAGTAAGTATAGTACCAAGAGAAATCCACACTGTGTGCC 180
DB 1901 AGCACTGTGACTTCTTTAAGTAAGTATAGTACCAAGAGAAATCCACACTGTGTGCC 1960

QY 181 TTTTACAAGCTATCAAGTCTCTGAAGATGACCCAACTTGCACAAAGAAATTTTGGCCA 240
DB 1961 TTTTACAAGCTATCAAGTCTCTGAAGATGACCCAACTTGCACAAAGAAATTTTGGCCA 2020

QY 241 CCAATTTGGATTCAGAGGCTCTCTTCTGACTATATCTCTCCGAGAAATTTCTCTTTG 300
DB 2021 CCAATTTGGATTCAGAGGCTCTCTTCTGACTATATCTCTCCGAGAAATTTCTCTTTG 2080

QY 301 AAGTACTACTGATTTACATTTGATGGATGCTCTACAGCTCATGATCAGCGTG 360
DB 2081 AAGTACTACTGATTTACATTTGATGGATGCTCTACAGCTCATGATCAGCGTG 2140

QY 361 GAAAGAAATATCTACTGCTGCTTCATATATATGTTGGTCTCCAGGGTCAATAGAAATG 420
DB 2141 GAAAGAAATATCTACTGCTGCTTCATATATATGTTGGTCTCCAGGGTCAATAGAAATG 2196

QY 421 ACGATCAGGTGGAGGACTCCCAATATCTAGCTTCTCGATATGATTTCAATGACTAGATC 480
DB 2197 ACGATCAGGTGGAGGACTCCCAATATCTAGCTTCTCGATATGATTTCAATGACTAGATC 2256

QY 481 GTGTGGGCATCCACGGCTGGTCTCTATGGAGGATACCTCTCCCTGATGGCAATTAATGAGA 540
DB 2257 GTGTGGGCATCCACGGCTGGTCTCTATGGAGGATACCTCTCCCTGATGGCAATTAATGAGA 2316

QY 541 GGTTCAGATATCTTCAGGGTTGCTATTTGCTGGGGCCCGAGTCACCTGTGGATCTTCATG 600
DB 2317 GGTTCAGATATCTTCAGGGTTGCTATTTGCTGGGGCCCGAGTCACCTGTGGATCTTCATG 2376

QY 601 ATACAGGATACCGAACCTTTATATGGGTTCACCTGACGAGAAATGAACAGGGCTATTACT 660
DB 2377 ATACAGGATACCGAACCTTTATATGGGTTCACCTGACGAGAAATGAACAGGGCTATTACT 2436

QY 661 TAGGATCTGGGCCATGCAAGCAAGAAAGTTCCCTCTCGAACCAAACTGTTTACTGCTCT 720
DB 2437 TAGGATCTGGGCCATGCAAGCAAGAAAGTTCCCTCTCGAACCAAACTGTTTACTGCTCT 2496

QY 721 TACATGTTTCTCTGGATGAGAAATGTCATTTTGCACATACCAGTATATTAATGAGTTTTT 780
DB 2497 TACATGTTTCTCTGGATGAGAAATGTCATTTTGCACATACCAGTATATTAATGAGTTTTT 2556

QY 781 TAGTGAGGCTGGAAGCCATATGATTTACAGATCTATCTCAGGAGAGACACAGCATAA 840
DB 2557 TAGTGAGGCTGGAAGCCATATGATTTACAGATCTATCTCAGGAGAGACACAGCATAA 2616

QY 841 GAGTTCCTGAATCGGAGAGAACATTAATGAATGCACTTTTGGCACTACTCTCAAGAAACC 900
DB 2617 GAGTTCCTGAATCGGAGAGAACATTAATGAATGCACTTTTGGCACTACTCTCAAGAAACC 2676

QY 901 TTGGATCAGATATGCTGCTCTAAAGTGATATAATTTTGACCTGTGTAGAACTCTCTGG 960
DB 2677 TTGGATCAGATATGCTGCTCTAAAGTGATATAATTTTGACCTGTGTGTAGAACTCTCTGG 2736

QY 961 TATACACTGGCTATTTTACCAATCAGGAGGTTTAAATCAACAGAAACACAGAAATTCATC 1020
DB 2737 TATACACTGGCTATTTTAAACCAATCAGGAGGTTTAAATCAACAGAAACACAGAAATTCATC 2796

QY 1021 ATCACATTTTGTATCTGCTGCAATGAAATCTACTCTCGAAATAAATGTGTGCATGCA 1080
DB 2797 ATCACATTTTGTATCTGCTGCAATGAAATCTACTCTCGAAATAAATGTGTGCATGCA 2856

QY 1081 GGGGTCTACGGTTTGTGGTAGTAATCTAATACCTTTAACCCCACTGTCTCAAAATCAAATG 1140
DB 2857 GGGGTCTACGGTTTGTGGTAGTAATCTAATACCTTTAACCCCACTGTCTCAAAATCAAATG 2916

QY 1141 ATACATATTTCTGAGAGAGCCAGCAATACCATAGAATTAATAAAAAA 1197
DB 2917 ATACATATTTCTGAGAGAGCCAGCAATACCATAGAATTAATAAAAAA 2973

RESULT 3
ADRA41222/c
ID ADRA41222 standard; cDNA; 2083 BP.
XX
AC ADRA41222;
XX
DT 07-OCT-2004 (first entry)
XX
DE Human CD-like molecule HQDES11 cDNA, seq ID NO:21.
XX
KW Human; CD-like molecule; cluster of differentiation; diagnosis;
KW prevention; immune disorder; immunodeficiency; autoimmune disorder;
KW blood-related disorder; haematological disorder; haemostatic disorder;
KW thrombolytic disorder; hyperproliferative disorder; cancer; tumour;
KW apoptotic disorder; cardiovascular disorder; respiratory disorder;
KW angiogenic disorder; neovascularisation; neurological disorder;
KW endocrine disorder; reproductive system disorder; infectious disease;
KW gastrointestinal disorder; drug screening; tissue regeneration;
KW chemotaxis; gene therapy; antibody therapy; drug targeting;
KW chromosome mapping; forensic analysis; immunophenotyping; cytostatic;
KW haemostatic; tranquiliser; vulnery; antiinflammatory; nephrotropic;
KW cardiant; anti-allergic; anti-HIV; anti-rheumatic; antiarthritic;
KW antipruritic; immunosuppressive; vasotropic; nootropic; neuroprotective;
KW antithyroid; thyromimetic; gynaecological; virucide; hepatotropic;
KW antibacterial; dermatological; chromosome 15q22; gene; ss.
OS Homo sapiens.
XX
FN WO200226930-A2.
XX
PD 04-APR-2002.
XX
PF 25-SEP-2001; 2001WO-US029838.
XX
PR 26-SEP-2000; 2000US-0235484P.
XX
PI (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Birse CE;
XX
DR WPI, 2002-405050/43.
XX
DR P-PSDB; ADRA41398.

PT Novel polynucleotides and polypeptides useful for treating, preventing or
PT ameliorating cardiovascular, renal, neurovascular, and autoimmune
XX disorders.
PS Claim 4; SEQ ID NO 21; 1243bp; English.
XX

CC The invention relates to 167 novel human CD (cluster of differentiation)-
CC like molecules (ADR41388-ADR41563) and to cDNAs encoding them (seqid11)-
XX
XX Sequence 2083 BP; 590 A; 440 C; 418 G; 635 T; 0 U; 0 Other;

Query Match 86.9%; Score 1040; DB 7; Length 2083;
Best Local Similarity 89.1%; Fred. No. S.1e-293;
Matches 1197; Conservative 0; Mismatches 0; Indels 147; Gaps 1;

Qy 1 ATTTGAGGACCAAGAGCTCCCTTTAGAGCATCACCTGTACGTAGTCAGTTAGCTAA 60
Db ATTTGAGGACCAAGAGCTCCCTTTAGAGCATCACCTGTACGTAGTCAGTTAGCTAA 1301

Qy 61 ATCTGGAGAGTGACAAGGCTGACTGACCGTGGCTACTCATCTTGTGTCATCAGTC 120
Db ATCTGGAGAGTGACAAGGCTGACTGACCGTGGCTACTCATCTTGTGTCATCAGTC 1241

Qy 121 AGCACTGTGACTCTTTTATAAGTATAGTAACTGACCAAGAAATCCACACTGTGTGCC 180
Db AGCACTGTGACTCTTTTATAAGTATAGTAACTGACCAAGAAATCCACACTGTGTGCC 1181

Qy 181 TTTTACAAGCTATCAAGTCTGGAAGATGACCAACTTGCAGAAATCCACACTGTGTGCC 240
Db TTTTACAAGCTATCAAGTCTGGAAGATGACCAACTTGCAGAAATCCACACTGTGTGCC 1121

Qy 241 CCATTTTGGATTGACAGGCTCTCTTCTGACTATATCTCTCCAGAAATTTCTCTTTG 300
Db CCATTTTGGATTGACAGGCTCTCTTCTGACTATATCTCTCCAGAAATTTCTCTTTG 1061

Qy 301 AAGTACTACTGGATTACATGTATGGATGCTTACAGAGCTCATGATCAGACCTG 360
Db AAGTACTACTGGATTACATGTATGGATGCTTACAGAGCTCATGATCAGACCTG 1001

Qy 361 GAAAGAAATATCTACTGCTGCTTCTATATATGTTGCTCTCA----- 403
Db GAAAGAAATATCTACTGCTGCTTCTATATATGTTGCTCTCA----- 941

Qy 404 ----- 403
Db ----- 881

Qy 404 ----- 403
Db ----- 881

Qy 880 TTGTAGTATGACAAACAGGGATCTGTCTACCGAGGGCTTAAATTTGAAGGCGCTTTA 821
Qy 404 -----GGTCAATAGAAATTTGACGATCAGGTGGAAGGACTCCAATATCTAGCTT 453
Db AATATAAATGGGTCAATAGAAATTTGACGATCAGGTGGAAGGACTCCAATATCTAGCTT 761

Qy 454 CTCGATATGATTTTCACTGACTTAGATCGTGTGGGCATCCACGGCTGGTCTATGGAGGAT 513
Db CTCGATATGATTTTCACTGACTTAGATCGTGTGGGCATCCACGGCTGGTCTATGGAGGAT 701

Qy 514 ACCTCTCCCTGATGGCATTAAATGACAGAGGTGAGATATCTTCAGGGTTGCTATTGCTGGG 573
Db ACCTCTCCCTGATGGCATTAAATGACAGAGGTGAGATATCTTCAGGGTTGCTATTGCTGGG 641

Qy 574 CCCAGTCACTCTGAGGATCTTCTATGATACAGATACACGGAACTGATATGGGTGAC 633
Db CCCAGTCACTCTGAGGATCTTCTATGATACAGATACACGGAACTGATATGGGTGAC 581

Qy 634 CTGACCAAGTAAACAGGCTTATCTAGGATCTGTGGCCATGCAAGCAAGAAATGTTCC 693
Db CTGACCAAGTAAACAGGCTTATCTAGGATCTGTGGCCATGCAAGCAAGAAATGTTCC 521

Qy 694 CCTCTGACCAAACTGTTTACTGCTCTTACATGTTTCTTGGATGAAATGTCATTTTG 753
CCTCTGACCAAACTGTTTACTGCTCTTACATGTTTCTTGGATGAAATGTCATTTTG

Db 520 CCTCTGAACCAAACTGTTTACTGCTCTTACATGGTTTCTCTGGATGAGAAATGTCATTTTG 461
Qy 754 CACATACCACTATATTACTGAGTTTTTTAGTGAGGGCTGGAAAGCCATATGATTTACAGA 813
Db 460 CACATACCACTATATTACTGAGTTTTTTAGTGAGGGCTGGAAAGCCATATGATTTACAGA 401
Qy 814 TCTATCTCAGGAGAGACACAGCATAAAGAGTTTCTGAAATCGGGAGAACATTTATGAACCTGC 873
Db 400 TCTATCTCAGGAGAGACACAGCATAAAGAGTTTCTGAAATCGGGAGAACATTTATGAACCTGC 341
Qy 874 ATCTTTTGCACCTACCTTCAAGAAAACCTTTGGATCACGTTATTTGCTCTCTAAAAGTGATAT 933
Db 340 ATCTTTTGCACCTACCTTCAAGAAAACCTTTGGATCACGTTATTTGCTCTCTAAAAGTGATAT 281
Qy 934 AATTTTGACCTGTGTAGAACTCTCTGGTATACACCTGGCTATTTAACCAAAATGAGGAGTT 993
Db 280 AATTTTGACCTGTGTAGAACTCTCTGGTATACACCTGGCTATTTAACCAAAATGAGGAGTT 221
Qy 994 TAATCAACAGAAAACACAGAAATTTGATCATCATTTTTCATACCTGCCCATGTAACATCTAC 1053
Db 220 TAATCAACAGAAAACACAGAAATTTGATCATCATTTTTCATACCTGCCCATGTAACATCTAC 161
Qy 1054 TCCTGAAAATAAATGTTGGTCCCATGCAGGGGTCTACGGTTTGTGGTAGTAACTAAATACC 1113
Db 160 TCCTGAAAATAAATGTTGGTCCCATGCAGGGGTCTACGGTTTGTGGTAGTAACTAAATACC 101
Qy 1114 TTAACCCCATGCTCAGAAATCAAAATGATATATTTCTGAGAGACCCAGCAATACCATTA 1173
Db 100 TTAACCCCATGCTCAGAAATCAAAATGATATATTTCTGAGAGACCCAGCAATACCATTA 41
Qy 1174 AGAATTACTAAAAAANAAAAA 1197
Db 40 AGAATTACTAAAAAANAAAAA 17

RESULT 4

ID ADL13374 standard; cDNA; 2797 BP.

XX AC ADL13374;

XX DT 06-MAY-2004 (first entry)

XX DE Human steroid-induced C3A liver cell cDNA #1103.

XX KW ss; gene; Hepatotropic; Gene therapy; Wilson disease; liver disorder;
KW steroid therapy; cirrhosis; hepatitis; human; C3A liver cell.

XX OS Homo sapiens.

XX FN US6673549-B1.

XX PD 06-JAN-2004.

XX PF 12-OCT-2001; 2001US-00976594.

XX PR 12-OCT-2000; 2000US-0240409P.

XX PA (INCY-) INCYTE CORP.

XX PI Furness LM, Buchbinder JL;

XX DR WPI; 2004-068610/07.

XX PT Combination useful for preparing a composition for treating liver
PT disorders associated with steroid therapy, e.g., cirrhosis or hepatitis,
PT comprises cDNAs that are differentially expressed in response to steroid
PT treatment.

XX PS Claim 1; SEQ ID NO 1103; 141pp; English.

XX CC The invention relates to a combination comprising cDNAs that are
CC differentially expressed in response to steroid treatment. Also included

are the following: a high throughput method for using a cDNA to detect differential expression of nucleic acids in a sample; and a high throughput method of screening molecules or compounds to identify a ligand that specifically binds a cDNA. The sample is from a subject with Wilson disease and comparison of a standard defines a stage of that disease. The high throughput method of screening molecules or compounds to identify a ligand that specifically binds a cDNA comprises: combining the combination with molecules or compounds under conditions to allow specific binding; and detecting specific binding between each cDNA and at least one molecule or compound. The molecules or compounds are regulatory proteins. The combination is useful for preparing a composition for treating liver disorders associated with steroid therapy, e.g., cirrhosis or hepatitis. The present sequence represents a human cDNA which is differentially expressed in steroid-induced C3A liver cells. Note: The sequence data for this patent did not form part of the printed specification but was obtained in electronic format directly from USPTO at seqdata.uspto.gov/sequence.html.

Sequence 2797 BP; 866 A; 551 C; 595 G; 785 T; 0 U; 0 Other;

Query Match 86.9%; Score 1040; DB 12; Length 2797;
Best Local Similarity 89.1%; Pred. No. 5.8e-293;
Matches 1197; Conservative 0; Mismatches 0; Indels 147; Gaps 1;
QY 1 ATTTTGAAGCCACCAAGACTCCCTTTAGACATCAGCTGACGTAGTACGTAA 60
DB 1452 ATTTTGAAGCCACCAAGACTCCCTTTAGACATCAGCTGACGTAGTACGTAA 1511
QY 61 ATCTGAGAGGTGACAGGCTGACGCGGTGCTACTCATTTCTTGTGCATCAGTC 120
DB 1512 ATCTGAGAGGTGACAGGCTGACGCGGTGCTACTCATTTCTTGTGCATCAGTC 1571
QY 121 AGCACTGTGACTCTTTATAAGTAAGTATAGTAACAGAGAATCCACACTGTGTGCC 180
DB 1572 AGCACTGTGACTCTTTATAAGTAAGTATAGTAACAGAGAATCCACACTGTGTGCC 1631
QY 181 TTTCAGAGCTATCAAGTCTGAAGATGACCAACTGTGCAAAACAAAGAAATTTGGGCCA 240
DB 1632 TTTCAGAGCTATCAAGTCTGAAGATGACCAACTGTGCAAAACAAAGAAATTTGGGCCA 1691
QY 241 CCATTTTGGATTCCAGAGGTCCTTCTGCTGACTATACCTCCAGAAATTTCTCTTTTG 300
DB 1692 CCATTTTGGATTCCAGAGGTCCTTCTGCTGACTATACCTCCAGAAATTTCTCTTTTG 1751
QY 301 AAAGTACTACTGGATTTACATTTGATGGATGCTCTACAAGCCTCATGATCTACAGCCTG 360
DB 1752 AAAGTACTACTGGATTTACATTTGATGGATGCTCTACAAGCCTCATGATCTACAGCCTG 1811
QY 361 GAAAGAAATATCCTACTGTGCTGTTCAATATATGTTGGTGCCTCA----- 403
DB 1812 GAAAGAAATATCCTACTGTGCTGTTCAATATATGTTGGTGCCTCAAGTGTGTAATA 1871
QY 404 ----- 403
DB 1872 ATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCTAGCTCTCTAGTTATGTGG 1931
QY 404 ----- 403
DB 1932 TTGTAGTGATAGACACAGGGGATCTGTGTCACCGAGGCTTAAATTTGAAGGCCCTTTA 1991
QY 404 -----GGGTCAAAATAGAAATTTGACGATCAGGTGGAAGGACTCCAATATCTAGCTT 453
DB 1992 AATATAAATGGGTCAAAATAGAAATTTGACGATCAGGTGGAAGGACTCCAATATCTAGCTT 2051
QY 454 CTCATATGATTTTCATGATCTAGATCGTGTGGGATCCAGGCTGGTCTCATGAGGAT 513
DB 2052 CTCGATATGATTTTCATGATCTAGATCGTGTGGGATCCAGGCTGGTCTCATGAGGAT 2111
QY 514 ACCTCTCCCTGATGGCATTAATGCAGAGGTGAGATATCTTCAGGGTGTCTATTTCTGGGG 573
DB 2112 ACCTCTCCCTGATGGCATTAATGCAGAGGTGAGATATCTTCAGGGTGTCTATTTCTGGGG 2171
QY 574 CCCAGTCACTGTGTGATCTTCTATGATACAGGATACACGGAAAGCTTATATGGGTCAAC 633

DB 2172 CCCAGTCACTGTGTGATCTTCTATGATACAGGATACACGGAAAGCTTATATGGGTCAAC 2231
QY 634 CTGACCAAGATTAACAGGGCTATTACTTAGGATCTGTGCCATGCAACAGAAAGTTCC 693
DB 2232 CTGACCAAGATTAACAGGGCTATTACTTAGGATCTGTGCCATGCAACAGAAAGTTCC 2291
QY 694 CCTCTGAACAAATCGTTTACTGCTCTTACATGCTTCTGATGAGATGTCATTTTG 753
DB 2292 CCTCTGAACAAATCGTTTACTGCTCTTACATGCTTCTGATGAGATGTCATTTTG 2351
QY 754 CACATACCACTATTACTGAGTGTGTTTATGAGGCTGGAAGCCATATGATTTACAGA 813
DB 2352 CACATACCACTATTACTGAGTGTGTTTATGAGGCTGGAAGCCATATGATTTACAGA 2411
QY 814 TCTATCCTCAGGAGACACAGCATAAAGATTCTCTGAATCGGGAACATTTATGAATGC 873
DB 2412 TCTATCCTCAGGAGACACAGCATAAAGATTCTCTGAATCGGGAACATTTATGAATGC 2471
QY 874 ATCTTTTGCATCTTCAAGAAACCTTGGATCAGTATTTGCTCTTAAAGTGATAT 933
DB 2472 ATCTTTTGCATCTTCAAGAAACCTTGGATCAGTATTTGCTCTTAAAGTGATAT 2531
QY 934 AATTTTGCACCTGTGTAAGTCTCTGTATACACTGGCTATTAAACCAAATGAGGAGTT 993
DB 2532 AATTTTGCACCTGTGTAAGTCTCTGTATACACTGGCTATTAAACCAAATGAGGAGTT 2591
QY 994 TAATCAACAGAAACACAGAAATGATCATCATATTTGATACCTGCCATGTAACATCTAC 1053
DB 2592 TAATCAACAGAAACACAGAAATGATCATCATATTTGATACCTGCCATGTAACATCTAC 2651
QY 1054 TCCTGAAAATAAATGTGGTGCATGAGGGGTCTAGGTTTGTGTAGTAACTTAATACC 1113
DB 2652 TCCTGAAAATAAATGTGGTGCATGAGGGGTCTAGGTTTGTGTAGTAACTTAATACC 2711
QY 1114 TTAACCCCACTGCTCAAAATCAAAATGATACATTTCTCTGAGAGACCAGCAATACCATA 1173
DB 2712 TTAACCCCACTGCTCAAAATCAAAATGATACATTTCTCTGAGAGACCAGCAATACCATA 2771
QY 1174 AGAATTACTAAAAAANAAAAA 1197
DB 2772 AGAATTACTAAAAAANAAAAA 2795
RESULT 5
AAH99934
ID AAH99934 standard; cDNA; 3143 BP.
XX AC AAH99934;
XX DT 12-APR-2002 (first entry)
XX DE cDNA encoding 21953 human prollyl oligopeptidase.
XX KW 21953 prollyl oligopeptidase; human; proline; endopeptidase; cancer;
cardiovascular disease; autoimmune disease; atopic allergy;
neural disorder; vascular disorder; prostate disorder; cytostatic;
antidiabetic; antiarthritic; antiasthmatic; antiinflammatory;
diabetes mellitus; arthritis; multiple sclerosis; asthma;
Grave's disease; neuronal disorder; demyelinating disease; ss.
OS Homo sapiens.
FH Key Location/Qualifiers
FT CDS 229..2877
FT /tag= a
FT /product= "21953 prollyl oligopeptidase"
FT /note= "This region is specifically claimed in claim 2"
XX WO200179473-A2.
XX 25-OCT-2001.
XX


```
QY 1174 AGAATTACTAAAAA 1193
DB 3101 AGAATTACTAAAAA 3120

RESULT 7
AAD38956
ID AAD38956 standard; cDNA; 3120 BP.
XX
AC AAD38956;
XX
DT 23-SEP-2002 (first entry)
XX
DE Human dipeptidyl peptidase 8 (DPP8) cDNA.
XX
KW Human; dipeptidyl peptidase; DPP; neoplasia; type II diabetes; cirrhosis;
KW autoimmunity; human immuno deficiency virus; HIV infection; cytostatic;
KW graft rejection; antidiabetic; antiinflammatory; immunosuppressive;
KW antiviral; enzyme; gene; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 214..2862
FT /tag= a
FT /product= "Human DPP8 protein"
XX
PN WO200234900-A1.
XX
PD 02-MAY-2002.
XX
PF 29-OCT-2001; 2001WO-AU001388.
XX
PR 27-OCT-2000; 2000AU-00001078.
XX
PA (UNSY ) UNIV SYDNEY.
XX
PI Abbott CA, Gorrell MD;
XX
DR WPI; 2002-454646/48.
XX
DR P-PSDB; AAE24170.
XX
PT New dipeptidyl peptidase (DPP) peptides, useful for screening inhibitors
PT of DPP catalytic activity, which may be employed to treat e.g. neoplasia,
PT type II diabetes, cirrhosis, autoimmunity, graft rejection and HIV
PT infection.
XX
XX Example; Fig 1; 91pp; English.
XX
CC The present invention relates to dipeptidyl peptidase (DPP) proteins and
CC polynucleotides encoding such proteins. The DPP peptides are useful for
CC screening inhibitors of DPP catalytic activity. The inhibitors are useful
CC for treating neoplasia, type II diabetes, cirrhosis, autoimmunity, graft
CC rejection and HIV (human immuno deficiency virus) infection. The present
CC sequence is human DPP8 cDNA
XX
SQ Sequence 3120 BP; 936 A; 637 C; 706 G; 841 T; 0 U; 0 Other;

Query Match 86.5%; Score 1036; DB 6; Length 3120;
Best Local Similarity 89.0%; Pred. No. 9e-292;
Matches 1193; Conservative 0; Mismatches 0; Indels 147; Gaps 1;

QY 1 ATTTGAAGGCCAACAAAGACTCCCTTTAGAGCATCACCTGTACGTAGTCAGTTACGTAA 60
DB 1781 ATTTTGAAGGCCAACAAAGACTCCCTTTAGAGCATCACCTGTACGTAGTCAGTTACGTAA 1840
XX
QY 61 ATCCCTGGAGAGGTGACAAAGGCTGACTGACCGTGGCTACTCACATTCTTGTGTCATCAGTC 120
DB 1841 ATCCCTGGAGAGGTGACAAAGGCTGACTGACCGTGGCTACTCACATTCTTGTGTCATCAGTC 1900
XX
QY 121 AGCACTGTGACTTCTTTATAAGTAAGTATAGTAACCAAGAGAAATCCACACTGTGTGCC 180
DB 1901 AGCACTGTGACTTCTTTATAAGTAAGTATAGTAACCAAGAGAAATCCACACTGTGTGCC 1960
```

```
QY 181 TTACAAGCTATCAAGTCTCTGAAGATGACCCAACTTGGCAAAACAAGGAATTTTGGGCCA 240
DB 1961 TTACAAGCTATCAAGTCTCTGAAGATGACCCAACTTGGCAAAACAAGGAATTTTGGGCCA 2020
XX
QY 241 CCATTTTGGATTCAGCAGGTCCTTCTCTGACTATATCTCTCCAGAAAATTTTCTCTTTTG 300
DB 2021 CCATTTTGGATTCAGCAGGTCCTTCTCTGACTATATCTCTCCAGAAAATTTTCTCTTTTG 2080
XX
QY 301 AAAGTACTACTGGATTTTACATTTGTATGGGATCTCTACAAGCCTCATGATCTACAGCCTG 360
DB 2081 AAAGTACTACTGGATTTTACATTTGTATGGGATCTCTACAAGCCTCATGATCTACAGCCTG 2140
XX
QY 361 GAAAGAAAATATCCTACTGTCTGTCTCATATATGGTGGTCTCTCA----- 403
DB 2141 GAAAGAAAATATCCTACTGTCTGTCTCATATATGGTGGTCTCTCAGGTGCAAGTTGGTGAATA 2200
XX
QY 404 ----- 403
DB 2201 ATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTAGCCTCTCTAGGTTATGTGG 2260
XX
QY 404 ----- 403
DB 2261 TTGTAGTGATAGACAACAGGGGATCTGTACCAGGGCTTAAATTTTGAAGGCGCTTTA 2320
XX
QY 404 -----GGGTCAAATAGAAAATTTGACGATCAGGTGSAAGGACTCCAATATCTAGCTT 453
DB 2321 AATATAAAATGGGTCAAATAGAAAATTTGACGATCAGGTGSAAGGACTCCAATATCTAGCTT 2380
XX
QY 454 CTCGATATGATTTTCATTGACTTAGATCGTGTGGGCATCCAGGCTGTCTATCGAGAT 513
DB 2381 CTCGATATGATTTTCATTGACTTAGATCGTGTGGGCATCCAGGCTGTCTATCGAGAT 2440
XX
QY 514 ACCTCTCCCTGATGGCAATTAATGACAGAGTCAGATATCTTCAGGTTGTCTATTCCTGGGG 573
DB 2441 ACCTCTCCCTGATGGCAATTAATGACAGAGTCAGATATCTTCAGGTTGTCTATTCCTGGGG 2500
XX
QY 574 CCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAACGTTTATATGGGTCAAC 633
DB 2501 CCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAACGTTTATATGGGTCAAC 2560
XX
QY 634 CTGACAGAAATGAACAGGCTATTAATCTAGGATCTGTGGCCATCAAGAGCAAAAGTTCC 693
DB 2561 CTGACAGAAATGAACAGGCTATTAATCTAGGATCTGTGGCCATCAAGAGCAAAAGTTCC 2620
XX
QY 694 CCTCTGAACCAAATCGTTTACTGCTCTTACATGGTTTCTCTGGATGAGAATGTCCATTTTG 753
DB 2621 CCTCTGAACCAAATCGTTTACTGCTCTTACATGGTTTCTCTGGATGAGAATGTCCATTTTG 2680
XX
QY 754 CACATACCAAGTATATTTACTGAGTTTCTTAGTGAGGCTGGAAAGCCATATGATTTACAGA 813
DB 2681 CACATACCAAGTATATTTACTGAGTTTCTTAGTGAGGCTGGAAAGCCATATGATTTACAGA 2740
XX
QY 814 TCTATCCTCAGGAGAGACACAGCATAAGAGTTCTCTGAATCGGAGAACATTTATGAACCTGC 873
DB 2741 TCTATCCTCAGGAGAGACACAGCATAAGAGTTCTCTGAATCGGAGAACATTTATGAACCTGC 2800
XX
QY 874 ATCTTTTGCACTACCTTCAAGAAAACCTTGGATCACGTTATTTGCTGCTCTAAAAGTGATAT 933
DB 2801 ATCTTTTGCACTACCTTCAAGAAAACCTTGGATCACGTTATTTGCTGCTCTAAAAGTGATAT 2860
XX
QY 934 AATTTTGGACTCTGTAGAACTCTCTGGTATACACTGCTGCTATTTAAACAAATGAGGAGTT 993
DB 2861 AATTTTGGACTCTGTAGAACTCTCTGGTATACACTGCTGCTATTTAAACAAATGAGGAGTT 2920
XX
QY 994 TAATCAACAGAAAACACAGAAATGATCATACATTTTGTATCTGCTGCTATACATCTAC 1053
DB 2921 TAATCAACAGAAAACACAGAAATGATCATACATTTTGTATCTGCTGCTATACATCTAC 2980
XX
QY 1054 TCTGAAATTAATCTGTGTGCATGCAGGGGTCTACCGTTTGTGGTGTAGTATCTTAATACC 1113
DB 2981 TCTGAAATTAATCTGTGTGCATGCAGGGGTCTACCGTTTGTGGTGTAGTATCTTAATACC 3040
```

QY 1114 TTAACCCACATGCTCAAAATCAATGATACATATTTCTGAGAGACCCAGCAATACCATA 1173
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Dy 3041 TTAACCCACATGCTCAAAATCAATGATACATATTTCTGAGAGACCCAGCAATACCATA 3100
QY 1174 AGAATTACTAAAAAATAAAA 1193
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Dy 3101 AGAATTACTAAAAAATAAAA 3120
RESULT 8
ID ABK83328 standard; cDNA; 4309 BP.
XX AC ABK83328;
XX AC
XX 12-AUG-2002 (first entry)
DE cDNA encoding human DPRP-1 splice variant #4.
XX Human; serine protease; dipeptidyl peptidase IV-related protein; DPRP;
KW DPPIV; infection; human immunodeficiency virus; HIV-1; HIV-2; pain;
KW diabetes; infertility; obesity; anorexia; Parkinson's disease; stroke;
KW heart failure; hypertension; urinary retention; osteoporosis; cancer;
KW ulcer; allergy; cancer; psychotic disorder; neurological disorder;
KW dyskinesia; reproductive disorder; inflammatory disorder;
KW metabolic disorder; gene; ss.
XX
XX Homo sapiens.
OS
XX
XX WO200231134-A2.
PN
XX 18-APR-2002.
PD
XX
XX 12-OCT-2001; 2001WO-US031874.
PF
XX
XX 12-OCT-2000; 2000US-0240117P.
PR
XX (PERR) FERRING BV.
PA
XX Qi S, Akinsanya KO, Riviere PJ, Junien J;
PI
XX WPI; 2002-444178/47.
DR
XX P-PSDB; ABd61597.
DR
XX
XX New dipeptidyl peptidase IV-related proteins and nucleic acids encoding
PT the proteins useful for treating e.g. fungal, bacterial, protozoan and
PT viral infections, cancers, allergies, neurological disorders, or pain.
XX
XX Disclosure; Page 67-68; 113pp; English.
XX
XX The present invention relates to the isolation of novel human serine
CC proteases referred to as dipeptidyl peptidase IV (DPPIV)-related proteins
CC (DPRP). The dipeptidyl peptidase IV-related proteins (DPRP) and nucleic
CC acids encoding them are useful for treating infections such as fungal,
CC bacterial, protozoan and viral infections, particularly infections caused
CC by human immunodeficiency virus (HIV-1 or HIV-2), pain, diabetes,
CC precocious puberty, infertility, obesity, anorexia, bulimia, Parkinson's
CC disease, acute heart failure, hypotension, hypertension, urinary
CC retention, osteoporosis, angina pectoris, stroke, ulcers, asthma,
CC allergies, cancers, migraine, vomiting, psychotic and neurological
CC disorders (e.g. anxiety, dementia, or schizophrenia), and dyskinesias.
CC These may also be used in discovering therapeutic agents for the
CC treatment of reproductive, inflammatory and metabolic disorders. ABK83322
CC -ABK83343 encode human DPRP proteins
XX
XX Sequence 4309 BP; 1304 A; 790 C; 907 G; 1308 T; 0 U; 0 Other;
SQ
Query Match 86.3%; Score 1033.2; DB 6; Length 4309;
Best Local Similarity 88.7%; Pred. No. 6.9e-291;
Matches 1194; Conservative 0; Mismatches 3; Indels 149; Gaps 1;
QY 1 ATTTTGAGGCGACCAAGACTCCCTTTAGAGCATCACCTGTCAGTACGTTACGTAA 60
|||

Db 1261 ATTTTGAGGCGACCAAGACTCCCTTTAGAGCATCACCTGTCAGTACGTTACGTAA 1320
QY 61 ATCTGTGAGAGGTGACAAAGCTGACGACCGTGGTACTACATTTCTGTGTCATCAGTC 120
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Dy 1321 ATCTGTGAGAGGTGACAAAGCTGACGACCGTGGTACTACATTTCTGTGTCATCAGTC 1380
QY 121 AGCACTGTGACTTCTTTTATAAGTATAGTAAGTAAACAGAGAAATCCACACTGTGTGTC 180
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Dy 1381 AGCACTGTGACTTCTTTTATAAGTATAGTAAGTAAACAGAGAAATCCACACTGTGTGTC 1440
QY 181 TTTACAAGCTATCAAGTCTCTGAAGTACCCAACTTGCAGAAATTTTGGGGCA 240
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Dy 1441 TTTACAAGCTATCAAGTCTCTGAAGTACCCAACTTGCAGAAATTTTGGGGCA 1500
QY 241 CCATTTTGGATTCCAGCAGGTCTCTTCTGACTATCTCTCCAGAAATTTTCTTTTG 300
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Dy 1501 CCATTTTGGATTCCAGCAGGTCTCTTCTGACTATCTCTCCAGAAATTTTCTTTTG 1560
QY 301 AAAGTACTACTGGATTACATTTGATGGGATCTCTCAAGCCTCATGATCTACAGCTG 360
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Dy 1561 AAAGTACTACTGGATTACATTTGATGGGATCTCTCAAGCCTCATGATCTACAGCTG 1620
QY 361 GAAAGAAATATCTACTGTCTGCTTCATATATGGTGGTCTC----- 402
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Dy 1621 GAAAGAAATATCTACTGTCTGCTTCATATATGGTGGTCTCCTCAGGTGTCAGTTGGTAA 1680
QY 403 ----- 402
Db 1681 TAATCGGTTTAAAGGAGTCAAGTATTTCCGTTTGAATACCCCTAGCCTCTCTAGGTTATGT 1740
QY 403 ----- 402
Db 1741 GGTGTAGTAGACACACAGGGGATCTGTCAAGGGGCTTAAATTTGAAGGCGCCTT 1800
QY 403 -----AGGGTCAATAGAAATTTGACCATCAGGTGGAGGACTCCCAATATCTAGC 451
Db 1801 TAAATATAAAATGGGTCAATAGAAATGACCATCAGGTGGAGGACTCCCAATATCTAGC 1860
QY 452 TTCTCGATATGATTTTCATTTGACTTAGATCGTGGGCTATCCAGGCTCGTCTATGAGG 511
Db 1861 TTCTCGATATGATTTTCATTTGACTTAGATCGTGGGCTATCCAGGCTCGTCTATGAGG 1920
QY 512 ATACCTCTCCCTGATGGCAATTAATGACAGGTCAGATATCTTCAGGGTGTCTATGCTGG 571
Db 1921 ATACCTCTCCCTGATGGCAATTAATGACAGGTCAGATATCTTCAGGGTGTCTATGCTGG 1980
QY 572 GGCCCGAGTCACCTCTGAGGATCTTCTATGATACAGATACACGGAACGTATATGGGTCA 631
Db 1981 GGCCCGAGTCACCTCTGAGGATCTTCTATGATACAGATACACGGAACGTATATGGGTCA 2040
QY 632 CCCTGACCAAGTGAACAGGGCTTATTTAGGATCTGTGGCCATGCAAGCAGAAAAAGTT 691
Db 2041 CCCTGACCAAGTGAACAGGGCTTATTTAGGATCTGTGGCCATGCAAGCAGAAAAAGTT 2100
QY 692 CCCCTCTGAACCAAAATCGTTTACTGCTTACATGGTTTCTTGGATGAGATGTCATTT 751
Db 2101 CCCCTCTGAACCAAAATCGTTTACTGCTCTTACATGGTTTCTTGGATGAGATGTCATTT 2160
QY 752 TGCACATACCAGTATATTTACTGAGTTTCTTGTAGGGCTGGAAAGCCATATGATTTACA 811
Db 2161 TGCACATACCAGTATATTTACTGAGTTTCTTGTAGGGCTGGAAAGCCATATGATTTACA 2220
QY 812 GATCTATCTCTCAGGAGACACAGCATATAGATTTCTGATCGGAGAACATTTATGAAT 871
Db 2221 GATCTATCTCTCAGGAGACACAGCATATAGATTTCTGATCGGAGAACATTTATGAAT 2280
QY 872 GCATCTTTTGCACCTACCTTCAAGAAACCTTCGGATCAGTATTTGCTGCTTAAAGTAT 931
Db 2281 GCATCTTTTGCACCTACCTTCAAGAAACCTTCGGATCAGTATTTGCTGCTTAAAGTAT 2340
QY 932 ATATTTTGGACCTGTGTAGAACTCTCTGGTATACACTGGCTATTTTAAACCAATGAGGAG 991
Db 2341 ATATTTTGGACCTGTGTAGAACTCTCTGGTATACACTGGCTATTTTAAACCAATGAGGAG 2400

QY 992 TTTAATCAACAGAAAACACAGAAATTGATCATCATATTTTGATACCTGCGATTAACATCT 1051
Db |||||||
QY 2401 TTTAATCAACAGAAAACACAGAAATTGATCATCATATTTTGATACCTGCGATTAACATCT 2460
Db |||||||
QY 1052 ACTCCTGMAAATAAATGTGGTGCATGAGGGGCTCTACGGTTTGTGGTACTGATTAATA 1111
Db |||||||
QY 2461 ACTCCTGMAAATAAATGTGGTGCATGAGGGGCTCTACGGTTTGTGGTACTGATTAATA 2520
Db |||||||
QY 1112 CCTTAACCCACATGCTCAAAATCAAAATGATACATATTCCTGAGAGACCCAGCAATACCA 1171
Db |||||||
QY 2521 CCTTAACCCACATGCTCAAAATCAAAATGATACATATTCCTGAGAGACCCAGCAATACCA 2580
Db |||||||
QY 1172 TAGAATTACTAATAAAAAAAAAA 1197
Db |||||||
QY 2581 TAGAATTACTAATAAAAAAAAAA 2606
Db |||||||

RESULT 9

ID ABK83327 standard; cDNA; 4829 BP.
XX AC ABK83327;
XX AC
XX DT 12-AUG-2002 (first entry)
XX DE cDNA encoding human DPRP-1 splice variant #3.
XX KW Human; serine protease; dipeptidyl peptidase IV-related protein; DPRP;
KW DPPIV; infection; human immunodeficiency virus; HIV-1; HIV-2; pain;
KW diabetes; infertility; obesity; anorexia; Parkinson's disease; stroke;
KW heart failure; hypertension; urinary retention; osteoporosis; cancer;
KW ulcer; allergy; cancer; psychotic disorder; neurological disorder;
KW dyskinesia; reproductive disorder; inflammatory disorder;
KW metabolic disorder; gene; ss.
XX OS Homo sapiens.
XX PN WO200231134-A2.
XX PD 18-APR-2002.
XX PF 12-OCT-2001; 2001WO-US031874.
XX PR 12-OCT-2000; 2000US-0240117P.
XX PA (FERR) FERRING BV.
XX PI Qi S, Akinsanya KO, Riviere PJ, Junien J;
XX WPI: 2002-444178/47.
XX P-PSDB; ABG61596.

XX New dipeptidyl peptidase IV-related proteins and nucleic acids encoding
PT the proteins, useful for treating e.g. fungal, bacterial, protozoan and
PT viral infections, cancers, allergies, neurological disorders, or pain.
XX PS Disclosure; Page 65-66; 113pp; English.
XX CC The present invention relates to the isolation of novel human serine
CC proteases referred to as dipeptidyl peptidase IV (DPPIV)-related proteins
CC (DPRP). The dipeptidyl peptidase IV-related proteins (DPRP) and nucleic
CC acids encoding them are useful for treating infections such as fungal,
CC bacterial, protozoan and viral infections, particularly infections caused
CC by human immunodeficiency virus (HIV-1 or HIV-2), pain, diabetes,
CC precocious puberty, infertility, obesity, anorexia, bulimia, Parkinson's
CC disease, acute heart failure, hypotension, hypertension, urinary
CC retention, osteoporosis, angina pectoris, stroke, ulcers, asthma,
CC allergies, cancers, migraine, vomiting, psychotic and neurological
CC disorders (e.g. anxiety, dementia, or schizophrenia) and dyskinesias.
CC These may also be used in discovering therapeutic agents for the
CC treatment of reproductive, inflammatory and metabolic disorders. ABK83322
CC -ABK83343 encode human DPRP proteins

XX SQ Sequence 4829 BP; 1466 A; 886 C; 1017 G; 1460 T; 0 U; 0 Other;
Query Match 86.3%; Score 1033.2; DB 6; Length 4829;
Best Local Similarity 88.7%; Pred. No. 7.3e-291;
Matches 1194; Conservative 0; Mismatches 3; Indels 149; Gaps 1;
QY 1 ATTTGAGGACCAAGAGATCCCTTTTAGAGCATCAGCTGTACGTAGTACGTAGTTACGTAA 60
Db |||||||
QY 1781 ATTTGAGGACCAAGAGATCCCTTTTAGAGCATCAGCTGTACGTAGTACGTAGTTACGTAA 1840
Db |||||||
QY 61 ATCTGAGAGGTGACAAGGCTGACTGACCGTGGTACTCACATTCTTGTGTCATCAGTC 120
Db |||||||
QY 1841 ATCTGAGAGGTGACAAGGCTGACTGACCGTGGTACTCACATTCTTGTGTCATCAGTC 1900
Db |||||||
QY 121 AGCACTGTGACTTCTTTTATAAGTATAGTAACCAAGAAATCCACACTGTGTGTCCTC 180
Db |||||||
QY 1901 AGCACTGTGACTTCTTTTATAAGTATAGTAACCAAGAAATCCACACTGTGTGTCCTC 1960
Db |||||||
QY 181 TTTACAAGCTATCAAGTCTGGAAGATGACCAACTTGCAGAAACAAAGAAATTTTGGGCCA 240
Db |||||||
QY 1961 TTTACAAGCTATCAAGTCTGGAAGATGACCAACTTGCAGAAACAAAGAAATTTTGGGCCA 2020
Db |||||||
QY 241 CCATTTTGGATTGACGAGGTCTCTCTCTGACTATATCTCTCAGAAATTTTCTCTTTTG 300
Db |||||||
QY 2021 CCATTTTGGATTGACGAGGTCTCTCTCTGACTATATCTCTCAGAAATTTTCTCTTTTG 2080
Db |||||||
QY 301 AAGTACTACTGGATTACATTTGATGGGATGCTCTACAAGCCTCATGATCTACAGCCTG 360
Db |||||||
QY 2081 AAGTACTACTGGATTACATTTGATGGGATGCTCTACAAGCCTCATGATCTACAGCCTG 2140
Db |||||||
QY 361 GAAAGAAATATCCTACTGCTGTGTTTCAATATGTTGTTCTCTC----- 402
Db |||||||
QY 2141 GAAAGAAATATCCTACTGCTGTGTTTCAATATGTTGTTCTCTCAGGTGCGTGTGAA 2200
Db |||||||
QY 403 ----- 402
Db |||||||
QY 2201 TAATCGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTAGCCTCTCTAGGTTATGT 2260
Db |||||||
QY 403 ----- 402
Db |||||||
QY 2261 GGTGTAGTGATAGACAACAGGGGATCCTGTCAACGAGGGCTTTAAATTTTGAAGCGCCTT 2320
Db |||||||
QY 403 -----AGGTCAAATAGAAATTTGACGATCAGGTGGAAGGACTCCAATATCTACG 451
Db |||||||
QY 2321 TAAATATAAAATGGGTCAAATAGAAATTTGACGATCAGGTGGAAGGACTCCAATATCTACG 2380
Db |||||||
QY 452 TTCTCGATATGATTTTCAATTTGACTTTAGATCGTGTGGCATCCACGGCTGCTCTATGGAGG 511
Db |||||||
QY 2381 TTCTCGATATGATTTTCAATTTGACTTTAGATCGTGTGGCATCCACGGCTGCTCTATGGAGG 2440
Db |||||||
QY 512 ATACCTCTCCCTGATGGCATTAAATGACAGAGTCAGATATCTTCAGGGTTGCTATTCCTGG 571
Db |||||||
QY 2441 ATACCTCTCCCTGATGGCATTAAATGACAGAGTCAGATATCTTCAGGGTTGCTATTCCTGG 2500
Db |||||||
QY 572 GSCCCGATCACTCTGTGGATCTCTATGATACAGGATACAGGAACGTTATATGGGTCA 631
Db |||||||
QY 2501 GSCCCGATCACTCTGTGGATCTCTATGATACAGGATACAGGAACGTTATATGGGTCA 2560
Db |||||||
QY 632 CCCTGACAGAGTAAACAGGGCTATTACTTAGATCTGTGGCATTCGAAGCAGAAAAGTT 691
Db |||||||
QY 2561 CCCTGACAGAGTAAACAGGGCTATTACTTAGATCTGTGGCATTCGAAGCAGAAAAGTT 2620
Db |||||||
QY 692 CCCTCTGAAACCAATCGTTTACTGCTCTTAATGTTTCTGTGATGAGATGTCATTT 751
Db |||||||
QY 2621 CCCTCTGAAACCAATCGTTTACTGCTCTTAATGTTTCTGTGATGAGATGTCATTT 2680
Db |||||||
QY 752 TGCACATACCAAGTATATCTAGTTTCTTGTAGGGCTGGAAGCCATATGATTTTACA 811
Db |||||||
QY 2691 TGCACATACCAAGTATATCTAGTTTCTTGTAGGGCTGGAAGCCATATGATTTTACA 2740
Db |||||||
QY 812 GATCTATCTCTCAGGAGAGACACAGCAATAAGAGTTCTCTGAATCGGAGAACCATTTGAACT 871
Db |||||||

Db 2741 GATCTATCTCAGGAGAGACACAGCATAAGAGTTCTTGAATCGGAGAACATTATGAAC 2800
Qy 872 GCATCTTTGACATCTCCTCAAGAAACCTTGATCAGTATTCGTCTCTAAAGTGAT 931
Db 2801 GCATCTTTGACATCTCCTCAAGAAACCTTGATCAGTATTCGTCTCTAAAGTGAT 2860
Qy 932 ATAATTTGACCTGTGTAGAACTCTCTGGTATACACTGGCTATTAAACCAATGAGGAG 991
Db 2861 ATAATTTGACCTGTGTAGAACTCTCTGGTATACACTGGCTATTAAACCAATGAGGAG 2920
Qy 992 TTTAATCAACAGAAAAACAGAAATTGATCATACATTTTGTATCTGCGCATGTAAACATCT 1051
Db 2921 TTTAATCAACAGAAAAACAGAAATTGATCATACATTTTGTATCTGCGCATGTAAACATCT 2980
Qy 1052 ACTCTGAAATAAATGTGTGCTCATGCGAGGGTCTACGGTTTGTGTAGTAATCTAATA 1111
Db 2981 ACTCTGAAATAAATGTGTGCTCATGCGAGGGTCTACGGTTTGTGTAGTAATCTAATA 3040
Qy 1112 CTTAACCACACATGCTCAAAATCAAAATGATACATATTCCTGAGAGACCCAGCAATACCA 1171
Db 3041 CTTAACCACACATGCTCAAAATCAAAATGATACATATTCCTGAGAGACCCAGCAATACCA 3100
Qy 1172 TAAGATTACTTAAAAAATAAAAAA 1197
Db 3101 TAAGATTACTTAAAAAATAAAAAA 3126

RESULT 10

ABK12892
ID ABK12892 standard; cDNA; 3106 BP.
AC ABK12892;
XX
DT 09-APR-2002 (first entry)
XX
DE Human protéase PRTS-9 cDNA sequence.
KW Human; protéase; PRTS; gastrointestinal; Crohn's disease; cancer;
KW cardiovascular; atherosclerosis; autoimmune disorder; dermatitis;
KW inflammatory disorder; acquired immunodeficiency syndrome; AIDS;
KW cell proliferative disorder; developmental disorder; epilepsy;
KW Duchenne muscular dystrophy; epithelial disorder; neurological disorder;
KW reproductive disorder; endometriosis; ss.
XX

OS Homo sapiens.

XX Key Location/Qualifiers
FH CDS 203...2851
FT /*tag= a
FT /product= "Human protéase PRTS-9"
XX

XX WO200198468-A2.

XX 27-DEC-2001.

XX 13-JUN-2001; 2001WO-US019178.

XX 16-JUN-2000; 2000US-0212336P.

XX 22-JUN-2000; 2000US-0213955P.

XX 29-JUN-2000; 2000US-0215396P.

XX 07-JUL-2000; 2000US-0216821P.

XX 14-JUL-2000; 2000US-0218946P.

XX (INCY-) INCYTE GENOMICS INC.

XX Yue H, Elliott VS, Gandhi AR, Lal P, Au-Young J, Tribouley CM;

PI Deleage AM, Baughn MR, Nguyen DB, Lee EA, Hafalia A, Khan FA;

PI Walia NK, Yao MG, Lu DM, Patterson C, Tang YT, Walsh RT;

PI Azinza Y, Lu Y, Ramkumar J, Xu Y, Reddy R, Das D, Kearney L;

PI Kallick DA;

DR WPI; 2002-090437/12.

DR P-PSDB; RAU74749.

XX
PT Twenty one human proteases (referred to as PRTS-1 to PRTS-21), useful in
PT the diagnosis, treatment and prevention of gastrointestinal (e.g.
PT gastritis), cardiovascular (e.g. atherosclerosis) and cell proliferative
PT (e.g. cancer) disorders.
XX
PS Claim 5; Page 166-167; 177pp; English.
PS
CC The present invention relates to twenty one new human proteases, referred
CC to as PRTS-1 to PRTS-21. The PRTS polynucleotides and polypeptides of the
CC invention are useful in the diagnosis, treatment and prevention of
CC gastrointestinal e.g. gastritis, esophageal carcinoma and Crohn's
CC disease, cardiovascular e.g. atherosclerosis, hypertension and myocardial
CC infarction, autoimmune/inflammatory e.g. acquired immunodeficiency
CC syndrome (AIDS), allergies and osteoarthritis, cell proliferative e.g.
CC cancer, developmental e.g. Duchenne and Becker muscular dystrophy,
CC epithelial e.g. dermatitis, neurological e.g. epilepsy and Alzheimer's
CC disease and reproductive e.g. infertility and endometriosis disorders.
CC Numerous other examples of each disorder are given in the specification.
CC The present nucleic acid sequence encodes the human protease PRTS-9
CC protein of the invention
XX
SQ Sequence 3106 BP; 928 A; 633 C; 704 G; 841 T; 0 U; 0 Other;

Query Match 86.3%; Score 1033; DB 6; Length 3106;
Best Local Similarity 89.0%; Pred. No. 6.8e-291;
Matches 1190; Conservative 0; Mismatches 0; Indels 147; Gaps 1;

Qy 1 ATTTTGAAGCACCACCAAGACTCCCTTTAGAGCATCACCTGTACGTAGTCAGTTACGTAA 60
Db 1770 ATTTTGAAGCACCACCAAGACTCCCTTTAGAGCATCACCTGTACGTAGTCAGTTACGTAA 1829
Qy 61 ATCTGTGAGAGGTGACAAAGGTGACCTGACCTGACCTGACCTGACCTGACCTGACCTG 120
Db 1830 ATCTGTGAGAGGTGACAAAGGTGACCTGACCTGACCTGACCTGACCTGACCTGACCTG 1889
Qy 121 AGCACTGTGACTTCTTTTATAAGTAAGTATAGTAACAGAGAAATCCACATGTGTGTCCC 180
Db 1890 AGCACTGTGACTTCTTTTATAAGTAAGTATAGTAACAGAGAAATCCACATGTGTGTCCC 1949
Qy 181 TTTTCAAGCTATCAAGTCTGTAAGTGAACCACTTGGCAAAACAAAGAAATTTTGGGCA 240
Db 1950 TTTTCAAGCTATCAAGTCTGTAAGTGAACCACTTGGCAAAACAAAGAAATTTTGGGCA 2009
Qy 241 CCAATTTGGATTGACGAGGTCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
Db 2010 CCAATTTGGATTGACGAGGTCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2069
Qy 301 AAAGTACTACTGGATTACATTTGATGGATGCTGTACAAAGCTCATGATCTACAGCCTG 360
Db 2070 AAAGTACTACTGGATTACATTTGATGGATGCTGTACAAAGCTCATGATCTACAGCCTG 2129
Qy 361 GAAAGAAATATCTTACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 403
Db 2130 GAAAGAAATATCTTACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2189
Qy 404 ----- 403
Db 2190 ATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCCTAGCCTCTCTAGTATTATGTGG 2249
Qy 404 ----- 403
Db 2250 TTGTAGTGATAGACACAGGGGATCTGTTCACCGAGGGCTTAAATTTGAAGCGGCTTTTA 2309
Qy 404 -----GGTCAATAGAAATTCAGATCAGGTGGAAGGACTCCCAATATCTAGCTT 453
Db 2310 AATAATAAATGGGTCAAATAGAAATTCAGATCAGGTGGAAGGACTCCCAATATCTAGCTT 2369
Qy 454 CTCGATGATGATTTTCATTGACTTGTGTTGGGATCCACCGCTGGTCTCTATGAGGAT 513
Db 2370 CTCGATGATGATTTTCATTGACTTGTGTTGGGATCCACCGCTGGTCTCTATGAGGAT 2429
Qy 514 ACCTCTCCCTGATGGCAATTAATCAGAGGTTCAGATATCTTCAGGGTGTGCTATTGTGCGG 573

Db 1896 AAAGTACTACTGGATTTACATTTGATGGGATGCTCTACAAGGCTCATGATCTACAGCGCTG 1955
Qy 361 GAAAGAAATATCCTACTGCTGCTCTCATATATGTTGGTCTCTCA----- 403
Db 1956 GAAAGAAATATCCTACTGCTGCTCTCATATATGTTGGTCTCTCAAGTGCGAGTTGGTGAATA 2015
Qy 404 ----- 403
Db 2016 ATCGGTTTAAAGGNGTCAAGTATTTCCGCTTGATACCTACCTGCTCTAGTTATGTGG 2075
Qy 404 ----- 403
Db 2076 TTGTAGTATAGACACAGGGGATCTGTACCGAGGGCTTAAATTTGAAGCGCGCTTTA 2135
Qy 404 -----GGGTCAAAATAGAAATTTGACGATCAGGTGGAAGGACTCCAATATCTAGCTT 453
Db 2136 AATATAAAATGGGTCAAAATAGAAATTTGACGATCAGGTGGAAGGACTCCAATATCTAGCTT 2195
Qy 454 CTCGATATGATTTTCATTGACTTAGATCGTGTGGGCATCCAGCGCTGCTCTATGGAGAT 513
Db 2196 CTCGATATGATTTTCATTGACTTAGATCGTGTGGGCATCCAGCGCTGCTCTATGGAGAT 2255
Qy 514 ACTCTCCCTGATGGGCAATTAATGCAGAGGTTCAGATATCTTCAGGGTTCCTATTGCTGGG 573
Db 2256 ACTCTCCCTGATGGGCAATTAATGCAGAGGTTCAGATATCTTCAGGGTTCCTATTGCTGGG 2315
Qy 574 CCCAGTCACTCTGTGGATCTCTATGATACAGGATACACGGAAGTTATATGGTCAAC 633
Db 2316 CCCAGTCACTCTGTGGATCTCTATGATACAGGATACACGGAAGTTATATGGTCAAC 2375
Qy 634 CTGACAGAAATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAGTTCC 693
Db 2376 CTGACAGAAATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAGTTCC 2435
Qy 694 CCTCTGAACCAAAATCGTTTACTGCTCTTACATGTTTCTGTGAGAGATGTCATTTTG 753
Db 2436 CCTCTGAACCAAAATCGTTTACTGCTCTTACATGTTTCTGTGAGAGATGTCATTTTG 2495
Qy 754 CACATACAGTATATTACTGAGTTTATTAGTGGGCTGGAAGCCATATGATTTACAGA 813
Db 2496 CACATACAGTATATTACTGAGTTTATTAGTGGGCTGGAAGCCATATGATTTACAGA 2555
Qy 814 TCTATCTCAGGAGACACAGCATAGAGTTCTGAAATCGGAGAACATTTCAACTGC 873
Db 2556 TCTATCTCAGGAGACACAGCATAGAGTTCTGAAATCGGAGAACATTTCAACTGC 2615
Qy 874 ATCTTTTGCACCTTCAAGAAACCTTTGGATCAGTATTGCTGCTTAAAGTGATAT 933
Db 2616 ATCTTTTGCACCTTCAAGAAACCTTTGGATCAGTATTGCTGCTTAAAGTGATAT 2675
Qy 934 AATTTTGACCTGTGTAGAACTCTCTGGTATACACTGGCTATTATTAACCAAAATGAGAGGTT 993
Db 2676 AATTTTGACCTGTGTAGAACTCTCTGGTATACACTGGCTATTATTAACCAAAATGAGAGGTT 2735
Qy 994 TAATCAACAGAAACACAGAAATGATCATCAATTTTGATACCTGCCATGTAACATCTAC 1053
Db 2736 TAATCAACAGAAACACAGAAATGATCATCAATTTTGATACCTGCCATGTAACATCTAC 2795
Qy 1054 TCTGAAAAATAATGTGGTCCATGACGGGGTCTACGGTTTGTGGTAGTAATCTAATACC 1113
Db 2796 TCTGAAAAATAATGTGGTCCATGACGGGGTCTACGGTTTGTGGTAGTAATCTAATACC 2855
Qy 1114 TTAACCCCATGCTGTAATTAATGATGATATTTCTGAGAGACCCAGCAATACCATTA 1173
Db 2856 TTAACCCCATGCTGTAATTAATGATGATATTTCTGAGAGACCCAGCAATACCATTA 2915
Qy 1174 AGAATTACTAAAA 1187
Db 2916 AGAATTACTAAAA 2929

RESULT 12

ABV76411
ID ABV76411 standard; cDNA; 1821 BP.
XX
AC ABV76411;
XX
DT 07-FEB-2003 (first entry)
XX
DE Dipeptidyl peptidase IV protein 27.5-encoding cDNA.
XX
KW Dipeptidyl peptidase IV protein 27.5; recombinant production;
KW Stenotrophomonas maltophilia dipeptidyl peptidase IV homologue;
KW gene therapy; diabetes; tumour; cancer; cytostatic; gene; ss.
XX
OS Unidentified.
XX
FH Key Location/Qualifiers
FT CDS 826..1578
FT /*tag= a
FT /product= "Dipeptidyl peptidase IV protein 27.5"
XX
PN CN1342768-A.
XX
PD 03-APR-2002.
XX
PF 12-SEP-2000; 2000CN-00125127.
XX
PR 12-SEP-2000; 2000CN-00125127.
XX
PA (BODE-) BODE GENE DEV CO LTD SHANGHAI.
XX
PI Mao Y, Xie Y;
XX
DR WPI: 2002-529809/57.
DR P-PSDB; ABB99949.
XX
PT New dipeptidyl peptidase IV protein 27.5 and encoding polynucleotide,
PT useful for treating diabetes and tumours.
XX
PS Claim 6; Page 24-25 (Disclosure); 32pp; Chinese.
XX
CC The invention relates to dipeptidyl peptidase IV protein 27.5 (ABB99949)
CC and nucleic acids encoding it (ABV76411). The protein has a molecular
CC weight of 27.5 kD and has 40% identity and 58% homology over a 227 amino
CC acid stretch with Stenotrophomonas maltophilia dipeptidyl peptidase IV
CC (Genbank accession number D83263). The invention also relates to a method
CC for the recombinant production of the protein, an antagonist of the
CC protein, and the use of the protein, gene and antagonist in therapeutic
CC applications. Dipeptidyl peptidase IV protein 27.5 can be used in the
CC treatment of a variety of diseases such as diabetes and tumours. The
CC present sequence represents cDNA encoding dipeptidyl peptidase IV protein
CC 27.5
XX
SQ Sequence 1821 BP; 533 A; 374 C; 391 G; 523 T; 0 U; 0 Other;
Query Match 85.3%; Score 1021; DB 6; Length 1821;
Best Local Similarity 88.9%; Pred. No; 1.7e-287;
Matches 1178; Conservative 0; Mismatches 0; Indels 147; Gaps 1;
Qy 1 ATTTTGAAGGCCCAAGAGACTCCCTTTAGAGCATCACCTGTAGTACGTACGTAA 60
Db 497 ATTTTGAAGGCCCAAGAGACTCCCTTTAGAGCATCACCTGTAGTACGTAA 556
Qy 61 ATCTGAGAGGTGACAAAGGCTGACGCGGTGGTACTCACATCTTGTGTCATCAGTC 120
Db 557 ATCTGAGAGGTGACAAAGGCTGACGCGGTGGTACTCACATCTTGTGTCATCAGTC 616
Qy 121 AGCACTGTGACTCTTTTATAGTAAGTATAGTAACCAAGAAATCCACACTGTGTGCC 180
Db 617 AGCACTGTGACTCTTTTATAGTAAGTATAGTAACCAAGAAATCCACACTGTGTGCC 676
Qy 181 TTTCACAGCTATCAAGTCTTGAAGATGACCCCACTTGCACAAAGAAATTTTGGGCA 240
Db 677 TTTCACAGCTATCAAGTCTTGAAGATGACCCCACTTGCACAAAGAAATTTTGGGCA 736

QY 241 CCATTTTGGATTACGAGGTCCTTTCTGACTATACCTCTCCAGAAATTTTCCTTTTG 300
DB |||||
QY 737 CCATTTTGGATTACGAGGTCCTTTCTGACTATACCTCTCCAGAAATTTTCCTTTTG 796
DB |||||
QY 301 AAAGTACTACTGGATTACATTTGATGGATGCTCTACAAGCCTCATGATCTACAGCCTG 360
DB |||||
QY 797 AAAGTACTACTGGATTACATTTGATGGATGCTCTACAAGCCTCATGATCTACAGCCTG 856
DB |||||
QY 361 GAAAGAAATATCTTACTGTGCTGTTTCATATATGGTGGTCTCA----- 403
DB GAAAGAAATATCTTACTGTGCTGTTTCATATATGGTGGTCTCAGGTGAGTTGGTGAATA 916
QY 404 ----- 403
DB 917 ATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTAGCCTCTCTAGTTATGTGG 976
QY 404 ----- 403
DB 977 TTGTAGTGATAGACACAGGGGATCCTGTACCGAGGGCTTAAATTTGAAGGCGCCTTTA 1036
QY 404 -----GGCTCAATAGAAATGACGATCAGGTGAGGAGCTCGAATATCTAGCTT 453
DB 1037 AATATAAATGGGTCAATAGAAATTTGACGATCAGGTGAGGAGCTCCATATCTAGCTT 1096
QY 454 CTCGATATGATTTCTATGACTTAGATCGTGTGGGATCCACGGCTGCTCATGGAGGAT 513
DB 1097 CTCGATATGATTTCTATGACTTAGATCGTGTGGGATCCACGGCTGCTCATGGAGGAT 1156
QY 514 ACTCTCCCTGATGGCATTAATGACAGAGTTCAGATATCTTCAGGGTTGCTATGCTGGGG 573
DB 1157 ACTCTCCCTGATGGCATTAATGACAGAGTTCAGATATCTTCAGGGTTGCTATGCTGGGG 1216
QY 574 CCCAGTCACTGTGTGATCTTCTATGATACAGGATACACGGAACTTATATGGGTCAAC 633
DB 1217 CCCAGTCACTGTGTGATCTTCTATGATACAGGATACACGGAACTTATATGGGTCAAC 1276
QY 634 CTGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAGTTCC 693
DB 1277 CTGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAGTTCC 1336
QY 694 CCTCTGAACCAAAATCGTTTACTGCTCTTACATGTTTCTCGATGAGAATGTCATTTTG 753
DB 1337 CCTCTGAACCAAAATCGTTTACTGCTCTTACATGTTTCTCGATGAGAATGTCATTTTG 1396
QY 754 CACATACAGTATATTACTGAGTGTGTTTGTAGTGGGCTGGAAGCCATATGATTTACAGA 813
DB 1397 CACATACAGTATATTACTGAGTGTGTTTGTAGTGGGCTGGAAGCCATATGATTTACAGA 1456
QY 814 TCTATCCTCAGGAGACACAGCATAGAGTTCCTGAAATCGGAGACATATGAACTGC 873
DB 1457 TCTATCCTCAGGAGACACAGCATAGAGTTCCTGAAATCGGAGACATATGAACTGC 1516
QY 874 ATCTTTGCACTACCTTCAAGAAACCTTGGATCAAGTATGCTGCTCTAAAGTGNAT 933
DB 1517 ATCTTTGCACTACCTTCAAGAAACCTTGGATCAAGTATGCTGCTCTAAAGTGNAT 1576
QY 934 AATTTTGACCTGTGTAGAACTCTCTGGTATACACTGGCTATTTAAACCAATGAGGAGTT 993
DB 1577 AATTTTGACCTGTGTAGAACTCTCTGGTATACACTGGCTATTTAAACCAATGAGGAGTT 1636
QY 994 TAATCAACAGAAACACAGAAATGATCATCATTTTGTATACCTGCCATGTAAACATCTAC 1053
DB 1637 TAATCAACAGAAACACAGAAATGATCATCATTTTGTATACCTGCCATGTAAACATCTAC 1696
QY 1054 TCTGAAATTAATATGGTGGCCATGACAGGGGTCTACGGTTTGTGGTAGTAACTTAATPACC 1113
DB 1697 TCTGAAATTAATATGGTGGCCATGACAGGGGTCTACGGTTTGTGGTAGTAACTTAATPACC 1756
QY 1114 TTAACCCACATGCTCAAAATCAAAATGATACATATTTCTTGAGAGACCCAGCAATACCATA 1173
DB 1757 TTAACCCACATGCTCAAAATCAAAATGATACATATTTCTTGAGAGACCCAGCAATACCATA 1816

QY 1174 AGAAT 1178
DB 1817 AGAAT 1821
RESULT 13
ABK83332
ID ABK83332 standard; cDNA; 4685 BP.
XX AC ABK83332;
XX DT 12-AUG-2002 (first entry)
XX cDNA encoding human DPRP-1 splice variant #8.
XX Human; serine protease; dipeptidyl peptidase IV-related protein; DPRP;
KW DPPIV; infection; human immunodeficiency virus; HIV-1; HIV-2; pain;
KW diabetes; infertility; obesity; anorexia; Parkinson's disease; stroke;
KW heart failure; hypertension; urinary retention; osteoporosis; cancer;
KW ulcer; allergy; cancer; psychotic disorder; neurological disorder;
KW dyskinesia; reproductive disorder; inflammatory disorder;
KW metabolic disorder; gene; ss.
XX OS Homo sapiens.
PN WO200231134-A2.
XX PN 18-APR-2002.
XX PF 12-OCT-2001; 2001WO-US031874.
XX PR 12-OCT-2000; 2000US-0240117P.
XX PA (FERR) FERRING BV.
XX PI Qi S, Akinsanya KO, Riviere PJ, Junien J;
XX WPI; 2002-444178/47.
XX P-PSDB; ABG61601.
XX New dipeptidyl peptidase IV-related proteins and nucleic acids encoding
PT the proteins, useful for treating e.g. fungal, bacterial, protozoan and
PT viral infections, cancers, allergies, neurological disorders, or pain.
XX Disclosure; Page 75-76; 113pp; English.
XX The present invention relates to the isolation of novel human serine
CC proteases referred to as dipeptidyl peptidase IV (DPPIV)-related proteins
CC (DPRP). The dipeptidyl peptidase IV-related proteins (DPRP) and nucleic
CC acids encoding them are useful for treating infections such as fungal
CC bacterial, protozoan and viral infections, particularly infections caused
CC by human immunodeficiency virus (HIV-1 or HIV-2), pain, diabetes,
CC precocious puberty, infertility, obesity, anorexia, bulimia, Parkinson's
CC disease, acute heart failure, hypotension, hypertension, urinary
CC retention, osteoporosis, angina pectoris, stroke, ulcers, asthma,
CC allergies, cancers, migraine, vomiting, psychotic and neurological
CC disorders (e.g. anxiety, dementia, or schizophrenia), and dyskinesias.
CC These may also be used in discovering therapeutic agents for the
CC treatment of reproductive, inflammatory and metabolic disorders. ABK83322
CC -ABK83343 encode human DPRP proteins
XX
SQ Sequence 4685 BP; 1430 A; 853 C; 951 G; 1411 T; 0 U; 0 Other;
Query Match 84.7%; Score 1013.6; DB 6; Length 4685;
Best Local Similarity 91.3%; Pred. No. 3.9e-285;
Matches 1098; Conservative 0; Mismatches 99; Indels 5; Gaps 2;
QY 1 ATTTTGAAGGCACCAAGAGCTCCCTTTTAGAGCATCACCTGTACGTAGTTCAGTTACGTA 60
DB 1781 ATTTTGAAGGCACCAAGAGCTCCCTTTTAGAGCATCACCTGTACGTAGTTCAGTTACGTA 1840
QY 61 ATCTCTGAGAGGTGACAAGGCTGACGCGGTGCTACTCACCATTCTTCTGCTGCATCAGTC 120
DB |||||

Db	1841	A	CTCTGAGAGGTGACAGAGCTGACGTGACGGTGGCTACTCACAATCTTCTGCTGCATCAGTC	1900
Qy	121	A	GCACTGTGACTTCTTTTATAGTAAGTATAGTAACAGAGAATCCACACTGTGTGTC	180
Db	1901	A	GCACTGTGACTTCTTTTATAGTAAGTATAGTAACAGAGAATCCACACTGTGTGTC	1960
Qy	181	T	TTTACAAGCTATCAAGTCTGAAGATGACCAACTTGCAGAAACAAAGGAATTTTGGGCCA	240
Db	1961	T	TTTACAAGCTATCAAGTCTGAAGATGACCAACTTGCAGAAACAAAGGAATTTTGGGCCA	2020
Qy	241	C	ATTTTGGATTTCAG---CAGGTCTCTTCTCTGACT-ATAGCTCTCCAGAAATTTTCTC	295
Db	2021	C	ATTTTGGATTTCAGTCTCTGAGTGTGTGTAATATCGTTTAAAGGAGTCAAGTA	2080
Qy	296	T	TTTGAAGTACTACTGGAATTTACATTTGATGGATGCTCTACAAGCCTCATGATCTACA	355
Db	2081	T	TTTCCGCTTGAATACCTAGCCTCTCTAGTTTATGGTTGTAGTGATAGACACAGGG	2140
Qy	356	G	CTTGGAAAGAAATATCTTACTGTCTGTTCATATATGGTGTCTCTCAGGTCAAAATAGA	415
Db	2141	A	TCCTGTACCGAGGGCTTAAATTTGAGGCGCTTTAAATATAAATGGGTCAAAATAGA	2200
Qy	416	A	ATTGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTTCAATGACTT	475
Db	2201	A	ATTGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTTCAATGACTT	2260
Qy	476	A	ATGCTGTGGGCATCCAGGCTGGTCTTATGAGGATACCTCTCCCTGATGGCATTAAT	535
Db	2261	A	ATGCTGTGGGCATCCAGGCTGGTCTTATGAGGATACCTCTCCCTGATGGCATTAAT	2320
Qy	536	G	CAGAGTCAGATATCTTCAAGGTTCATATGCTGGGGCCCGAGTCACTCTGTGGATCTT	595
Db	2321	G	CAGAGTCAGATATCTTCAAGGTTCATATGCTGGGGCCCGAGTCACTCTGTGGATCTT	2380
Qy	596	C	TATGATACAGGATACACGAGCTTATATGGTCACTCAGCAGATGAAACAGGCTA	655
Db	2381	C	TATGATACAGGATACACGAGCTTATATGGTCACTCAGCAGATGAAACAGGCTA	2440
Qy	656	T	TACTTAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTGAACCAATCTTTTACT	715
Db	2441	T	TACTTAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTGAACCAATCTTTTACT	2500
Qy	716	G	CTCTTACATGGTTTCTTGATGAGAAATCTCATTTTGCATACCACTATATTACTGAG	775
Db	2501	G	CTCTTACATGGTTTCTTGATGAGAAATCTCATTTTGCATACCACTATATTACTGAG	2560
Qy	776	T	TTTTTACTGAGGGTGGAAAGCCATATGATTTACAGATCTATCTCAGGAGACACAG	835
Db	2561	T	TTTTTACTGAGGGTGGAAAGCCATATGATTTACAGATCTATCTCAGGAGACACAG	2620
Qy	836	C	ATAAGAGTTCTGGAATCGGAGAACATTTATGAATGCAATCTTTTGGCACTACCTTCAAGA	895
Db	2621	C	ATAAGAGTTCTGGAATCGGAGAACATTTATGAATGCAATCTTTTGGCACTACCTTCAAGA	2680
Qy	896	A	AAACCTTGGATCAGTATTTGCTCTCTAAAAGTGATATAATTTTGACCTGTGTAGAACTC	955
Db	2681	A	AAACCTTGGATCAGTATTTGCTCTCTAAAAGTGATATAATTTTGACCTGTGTAGAACTC	2740
Qy	956	T	CTGGTATACATGGCTATTTAACCAATAGAGAGGTTTATCAACAGAAACACAGAAAT	1015
Db	2741	T	CTGGTATACATGGCTATTTAACCAATAGAGAGGTTTATCAACAGAAACACAGAAAT	2800
Qy	1016	T	GATCATCATATTTGATACCTGCCATGTAACATCTCTCTGAAAATFAAATGTGGTCC	1075
Db	2801	T	GATCATCATATTTGATACCTGCCATGTAACATCTCTCTGAAAATFAAATGTGGTCC	2860
Qy	1076	A	TGAGGGGTCTACGGTTTGTGTAGTAATCTAATACCTTAAACCCCATGCTCAAAATC	1135
Db	2861	A	TGAGGGGTCTACGGTTTGTGTAGTAATCTAATACCTTAAACCCCATGCTCAAAATC	2920
Qy	1136	A	ATATGATACATATTTCTGAGAGACCCAGCAATACCATGAAATTTCTAAAAAATAAAAA	1195
Db	2921	A	ATATGATACATATTTCTGAGAGACCCAGCAATACCATGAAATTTCTAAAAAATAAAAA	2980

Qy	1196	AA	1197
Db	2981	AA	2982
RESULT 14			
ACA92421			
ID	ACA92421	standard; DNA; 2952	BP.
XX	ACA92421;		
AC	ACA92421;		
XX	15-JUL-2003	(first entry)	
DT	15-JUL-2003	(first entry)	
DE	DNA	encoding human PMM-6.	
XX	Human; protein	modification and maintenance molecule; PMM; cancer;	
KW	cell proliferation	disorder; atherosclerosis; neurological disorder;	
KW	epilepsy; Huntington's	disease; stroke; immune disorder; allergy;	
KW	inflammatory disorder; AIDS;	developmental disorder; hypothyroidism;	
KW	Cushing's syndrome; gastrointestinal	disorder; epithelial disorder;	
KW	infection; cytostatic; antiarteriosclerotic;	anticonvulsant; nootropic;	
KW	neuroprotective; cerebroprotective;	anti-HIV; antiallergic; vulnery;	
KW	antiinflammatory; thyromimetic;	gene; ds.	
XX	Homo sapiens.		
OS	WO2003031939-A2.		
PN	17-APR-2003.		
XX	11-OCT-2002; 2002WO-US032850.		
PF	12-OCT-2001; 2001US-0329689P.		
PR	25-OCT-2001; 2001US-0335703P.		
PR	09-NOV-2001; 2001US-0348887P.		
PR	28-NOV-2001; 2001US-0334145P.		
PR	06-DEC-2001; 2001US-0337451P.		
PR	14-DEC-2001; 2001US-0340584P.		
XX	(INCY-) INCYTE GENOMICS INC.		
XX	Ramkumar J, Gorvad AE, Baughn MR, Emerling BM, Yang J, Lee SY;		
PI	Tran UK, Becha SD, Duggan BM, Lee EA, Griffin JA, Li JK;		
PI	Sprague JW, Hafalia AJA, Chawla NK, Lehr-Mason PM, Kable AE, Yue H;		
PI	Marquis JP, Yao MG, Richardson TW, Tang TV, Jin P, Chien D;		
PI	Bhatia U, Burrill JD, Lee S, Blake JJ, Ho A, Zheng W;		
XX	WPI; 2003-430274/40.		
DR	P-PSDB; ABU92026.		
XX	New human protein modification and maintenance molecules (PMM), useful		
PT	for diagnosing, treating and preventing diseases or conditions associated		
PT	with the aberrant PMM expression e.g. cancer, atherosclerosis, or		
PT	infections.		
PS	Claim 5; Page 285-286; 311pp; English.		
XX	The present invention relates to the isolation of human protein		
CC	modification and maintenance molecules (PMM), and the polynucleotide		
CC	sequences encoding them. A total of 40 PMM polypeptides (designated PMM		
CC	-1 to PMM-40) are disclosed. The sequences of the invention are useful		
CC	for diagnosing a condition or disease associated with the expression of		
CC	PMM in a subject, preparing a polyclonal or monoclonal antibody, and		
CC	generating an expression profile of a sample containing the		
CC	polynucleotides. The diseases or conditions associated with decreased		
CC	expression or overexpression of PMM are cell proliferation disorders		
CC	(e.g. cancer, atherosclerosis), neurological disorders (e.g. epilepsy,		
CC	Huntington's disease, stroke), immune/inflammatory disorders, (e.g. AIDS,		
CC	allergies), developmental disorders (e.g. hypothyroidism, Cushing's		
CC	syndrome), gastrointestinal or epithelial disorders, and infections. The		
CC	PMM polypeptides or their fragments are useful in screening compounds		
CC	for effectiveness as agonists or antagonists of the polypeptides, or in		

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 4, 2006, 02:59:51 ; Search time 4483.98 Seconds
(without alignments)
12489.839 Million cell updates/sec

Title: US-10-825-632-4

Perfect score: 1197

Sequence: 1 atttgaaggcaccacaaagac.....ttactaaaaaaaaaaaaa 1197

Scoring table: IDENTITY NUC
Gap 10.0, Gapext 1.0

Searched: 41078325 segs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_hcc:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_gss1:*
10: gb_gss2:*
11: gb_gss3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	929.6	77.7	1292	4 AF175225	AF175225 Homo sapi
2	900	75.2	1265	4 AF176779	AF176779 Homo sapi
3	832	69.5	1278	4 AF173382	AF173382 Homo sapi
4	788	65.8	2292	4 CR609512	CR609512 full-leng
5	768	64.2	864	1 AL542617	AL542617 AL542617
6	766.2	64.0	804	5 BU687687	BU687687 UI-CF-EC1
7	758.4	63.4	2649	10 AY411615	AY411615 Homo sapi
8	726.4	60.7	760	1 AI917735	AI917735 tclld07.x
9	717.4	59.9	2649	10 AY411616	AY411616 Pan trogl
10	704.8	58.9	5517	4 AK029788	AK029788 Mus muscu
11	684.8	57.2	3143	4 AK016546	AK016546 Mus muscu
12	677	56.6	735	1 AL043338	AL043338 DKFp0340
13	674.2	56.3	910	2 BG479035	BG479035 60252611
14	673.8	56.3	708	5 BU622228	BU622228 UI-H-FH1
15	668.8	55.9	673	1 AW303607	AW303607 xv21b12.x
16	642	53.6	722	5 BU631054	BU631054 UI-H-FE1
17	633	52.9	634	1 AI819365	AI819365 wg61c02.x
18	625.4	52.2	627	2 BG752328	BG752328 602730802
19	599.2	50.1	1002	2 BE73691	BE73691 601569288
20	598.8	50.0	855	2 BI084885	BI084885 602869453
21	591.4	49.4	597	5 BU784521	BU784521 in16d12.x
22	586.4	49.0	2634	10 AY411617	AY411617 Mus muscu

C	23	578	48.3	639	1	AW958665	EST370735
C	24	577.2	48.2	1042	3	BM557438	AGENCOURT
C	25	574.4	48.0	577	1	AI243857	qh89f09.x
C	26	556.6	46.5	738	7	CN427234	170006000
C	27	551.4	46.1	590	3	BI711119	id90q08.x
C	28	549	45.9	804	7	CN438529	BE0401681
C	29	533.4	44.6	776	6	CB522171	UI-M-GH0-
C	30	532.4	44.5	905	6	CB209486	AGENCOURT
C	31	529.4	44.2	536	3	BI710784	id90g08.y
C	32	519.2	43.4	719	8	DR768184	ILLUMINIGEN
C	33	509.8	42.6	746	5	BY764370	BY764370
C	34	502.8	42.0	800	2	BG921267	BG921267
C	35	496.8	41.5	500	1	AA496257	aa24b08.s
C	36	493.6	41.2	520	1	AW469326	hc81h10.x
C	37	490	40.9	617	2	BE789604	601481552
C	38	483.6	40.4	970	2	BG403555	602419260
C	39	483	40.4	494	1	AA278626	zs78g11.s
C	40	482.2	40.3	797	6	CA749361	UI-M-FY0-
C	41	482	40.3	493	1	AA278625	zs78g11.f
C	42	476.2	39.8	1038	1	AL582206	AL582206
C	43	474.4	39.6	820	6	CF532140	UI-M-GH0-
C	44	473.2	39.5	554	1	AW469635	hd32b10.x
C	45	468.2	39.1	835	7	CV557345	UI-M-H20-

ALIGNMENTS

RESULT 1
AF175225 1292 bp mRNA linear HTC 01-AUG-2003
LOCUS Homo sapiens tissue-type aorta MSTP135 mRNA, complete cds.
DEFINITION AF175225
ACCESSION AF175225
VERSION AF175225.1 GI:33338055
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 1292)
AUTHORS Zhao, B., Xu, H.S., Tong, Y.K., Sheng, H., Qin, B.M., Liu, Y.Q., Liu, B.,
Wang, X.Y., Zhang, Q., Song, L., Gao, Y., Zhang, C.L., Ye, J., Ji, X.J.,
Liu, B.H., Lu, H., Chen, J.Z., Cai, M.Q., Zhang, W.Y., Teng, C.Y.,
Liu, Q., Yu, L.T., Lin, J., Gong, Q., Zhang, A.M., Gao, R.L. and Hui, R.T.
Direct Submission
TITLE Submitted (04-AUG-1999) Molecular Medicine Center for
JOURNAL Cardiovascular Disease, Cardiovascular Institute, CAMS & PUMC, 167,
Bei Li Shi Lu, Beijing 100037, P.R. China
FEATURES
source 1. 1292
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="aorta"
142..1023
/codon_start=1
/product="MSTP135"
/protein_id="AA013650.1"
/db_xref="GI:33338056"
/translation="MTQLAKQRNFGPPFIQVQLFTLLQKFSLLKVLKLLDLCMGCS
TSRPGTGLERNILLXLYGGPKVQLVNNRFGKVFRLNTLPGLVYVVGIDNRGS
CPRLKPEGAFFKMGQIETIDQVGLQVLAQRDFIDLRVGHGYSYGYSLMAL
MORSIDIPRVAIAGAPVTLWTFDTGYTERYMGHPDQEQGYVLGSVMAQSKFSEPN
RLLLLHGFLENVHFAHTSILLSFLVRAGKPYDLQIYPQERHSIRVPESGHEYLHLL
HYLQENLGSRIALKVI"

ORIGIN

Query Match 77.7%; Score 929.6; DB 4; Length 1292;
Best Local Similarity 87.1%; Pred. No. 4e-246;
Matches 1119; Conservative 0; Mismatches 15; Indels 151; Gaps 3;
64 CTGGAGAGGTGACAAAGCTGACTGACCGTGGCTACTCATTCTTGTGTCATGCTAGC 123

Db 1 CTGAGAGGTGACAGAGGCTGACTGACCGTGGCTACTCACATCTTGCTGCTCAGTCAGC 60
Qy 124 ACTGTGACTTCTTTTAAGTAAGTATAGTAAACAGAGAAATCCAACATGTGTGCTCCCTTT 183
Db 61 ACTGTGACTTCTTTTAAGTAAGTATAGTAAACAGAGAAATCCAACATGTGTGCTCCCTTT 120
Qy 184 ACAAGCTATCAAGTCTCTGAAGATGACCCCACTTGCACAAACAAAGAAATTTTGGGCCACCA 243
Db 121 ACAAGCTATCAAGTCTCTGAAGATGACCCCACTTGCACAAACAAAGAAATTTTGGGCCACCA 180
Qy 244 TTTTGGATTCAGCAGGTCCTCTTCTGCTGACTATATCTCTCCAGAAATTTTCTTTTGA 303
Db 181 TTTTGGATTCAGCAGGTCCTCTTCTGCTGACTATATCTCTCCAGAAATTTTCTTTTGA 240
Qy 304 GTACTACTGATTTACATTTGATGGGATGCTCTCAAGCCCTCATGATCTACA---GCCTG 360
Db 241 GTACTACTGATTTACATTTGATGGGATGCTCTCAAGCCCTCATGATCTACAGGCCG 300
Qy 361 GAAAGAAATATCTACTGT- GCTGTTTCATATATGTTGCTCTCA- ----- 403
Db 301 GAAAGAAATATCTACTGNGGCTGTTTCATATATGGGCCCTTAAGGTGCAGTTGGTGAAT 360
Qy 404 ----- 403
Db 361 AATCGGTTTAAAGAGTCAAGTTTTCGCTTGATATACCTTACCCTTTTAGTTATGTG 420
Qy 404 ----- 403
Db 421 GTTGNAGGATAGACAAACAGGGATCTGTCCCGAGGGCTTAATTTGAAGGCCCTTT 480
Qy 404 -----GGTCAATAGAAATGACATGAGTGAAGGACTCCATATCTAGCT 452
Db 481 AAATTTAAATGGGTCAATAGAAATGACATGAGTGAAGGACTCCATATCTAGCT 540
Qy 453 TCTCGATATGATTTCACTTGTAGATCGTGTGGGCATCCACGGCTGGTCTATGAGGGA 512
Db 541 TCTCGATATGATTTCACTTGTAGATCGTGTGGGCATCCACGGCTGGTCTATGAGGGA 600
Qy 513 TACCTCTCCCTGATGCTTAATGACAGAGGTGAGATATCTTCAGGGTGTCTATGCTGGG 572
Db 601 TACCTCTCCCTGATGCTTAATGACAGAGGTGAGATATCTTCAGGGTGTCTATGCTGGG 660
Qy 573 GCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACAGGAAACGTTATATGGGTCA 632
Db 661 GCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACAGGAAACGTTATATGGGTCA 720
Qy 633 CCTGACAGAAATGAACAGGCTATTACTTATGATCTGTGGCATGCAAGCAAGAAAGTTC 692
Db 721 CCTGACAGAAATGAACAGGCTATTACTTATGATCTGTGGCATGCAAGCAAGAAAGTTC 780
Qy 693 CCCTCTGAACCAATCGTTTACTGCTCTTACATGTTTCTCGGATGAGATGTCCATTTT 752
Db 781 CCCTCTGAACCAATCGTTTACTGCTCTTACATGTTTCTCGGATGAGATGTCCATTTT 840
Qy 753 GCATATACCAAGTATATTACTGATTTTATGAGGCTGGAAAGCCATATGATTTACAG 812
Db 841 GCATATACCAAGTATATTACTGATTTTATGAGGCTGGAAAGCCATATGATTTACAG 900
Qy 813 ATCTATCTCAGGAGACACAGCATTAAGTTCCTGAATCGGGACACATTAAGTCACTG 872
Db 901 ATCTATCTCAGGAGACACAGCATTAAGTTCCTGAATCGGGACACATTAAGTCACTG 960
Qy 873 CATCTTTGCACTACCTTCAAGAAACCTTGATCAGTATTCCTGCTCTTAAAGTGATA 932
Db 961 CATCTTTGCACTACCTTCAAGAAACCTTGATCAGTATTCCTGCTCTTAAAGTGATA 1020
Qy 933 TAATTTTGAACCTGTAGAACTCTCTGGTATACACTGGCTATTAAACAAATGAGGAGGT 992
Db 1021 TAATTTTGAACCTGTAGAACTCTCTGGTATACACTGGCTATTAAACAAATGAGGAGGT 1080
Qy 993 TTAATCAACAGAAACACAGAAATTGATCATCACTTTTGATACCTGCGCATGTAACTCTA 1052

Db 1081 TTAATCAACAGAAACACAGAAATTGATCATCACATTTTGATACCTGCCATGTAACTCTA 1140
Qy 1053 CTCTGTAATAATAATGCTGTCATGCAAGGGTCTAGGGTTTGAGTAACTAATCTAATAC 1112
Db 1141 CTCTGTAATAATAATGCTGTCATGCAAGGGTCTAGGGTTTGAGTAACTAATCTAATAC 1200
Qy 1113 CTTAACCCACATGCTCAAAATCAAAATGATACATATTCCTGAGAGACCCAGCATCAT 1172
Db 1201 CTTAACCCACATGCTCAAAATCAAAATGATACATATTCCTGAGAGACCCAGCATCAT 1260
Qy 1173 AAGAATTTACTAAAAA 1197
Db 1261 AAGAATTTCTAAAAA 1285

RESULT 2
AF176779
LOCUS Homo sapiens MSTP141 mRNA, complete cds. linear HTC 01-AUG-2003
DEFINITION AF176779
ACCESSION AF176779
VERSION AF176779.1 GI:33338069
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
REFERENCE 1 (bases 1 to 1265)
AUTHORS Hui, R. T., Liu, Y. Q., Wang, X. Y., Qin, B. M. and Sheng, H.
TITLE Homo sapiens normal aorta mRNA MST141
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1265)
AUTHORS Hui, R. T., Liu, Y. Q., Wang, X. Y., Qin, B. M. and Sheng, H.
TITLE Direct Submission
JOURNAL Submitted (10-AUG-1999) Molecular Medicine Center for
Cardiovascular Disease, Cardiovascular Institute, CAMS & PUMC, 167,
Bei Li Shi Lu, Beijing, 100037, P.R. China

FEATURES
source
1..1265
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="aorta"
268..741
/codon_start=1
/product="MSTP141"
/protein_id="AAQ13657.1"
/db_xref="GI:33338070"
/translation="MLYKPHDLQPGKYPVTLFIYGGPQVOLVNNRFKGVKYPRLNTL
ASLGVVVVVIDNRGSGHRGLKFGKYMGIIDDOEGLOVLAGRYDFIDLDVRG
IHGWSYGYLSLMLMQRSDIFRVAIAGAPVTLWIFYDTGYTERYMGHPQNEQG"

ORIGIN
Query Match 75.2%; Score 900; DB 4; Length 1265;
Best Local Similarity 86.2%; Pred. No. 6.5e-238;
Matches 1105; Conservative 0; Mismatches 5; Indels 172; Gaps 3;
Qy 64 CTGAGAGGTGACAAAGG-CTGACTGACCGTGGCTACTCACATTTCTTGCTGATCAGTCAG 122
Db 1 CTGAGAGGTGACAAAGGCTGACTGACCGTGGCTACTCACATTTCTTGCTGATCAGTCAG 60
Qy 123 CACTGTGACTCTTTTATAGTAAGTATAGTAACCAAGAAATCCACACTGTGTGCCCTT 182
Db 61 CACTGTGACTCTTTTATAGTAAGTATAGTAACCAAGAAATCCACACTGTGTGCCCTT 120
Qy 183 TAGAAGTATCAAGTCTCAAGATGACCCCACTTCACAAACAAAGAAATTTTGGGCCACC 242
Db 121 TAGAAGTATCAAGTCTCAAGATGACCCCACTTCACAAACAAAGAAATTTTGGGCCACC 180
Qy 243 ATTTTGGATTCAGCAGGTCCTCTTCTGACTACTCTCCAGAAATTTTCTCTTTTGA 302
Db 181 ATTTTGGATTCAGCAGGTCCTCTTCTGACTACTCTCTCCAGAAATTTTCTCTTTTGA 240


```
QY 303 AATACTACTCGATTTACATTTGATGGGATGCTCTCAAGACCTCATGATCTACAGCCTGGA 362
Db 241 AGTACTACTCGATTTACATTTGATGGGATGCTCTCAAGACCTCATGATCTACAGCCTGGA 300
QY 363 AAGAAATATCTACTGCTGCTTTTATATATGTTGGTCTCTCA----- 403
Db 301 AAGAAATATCTACTGCTGCTTTTATATATGTTGGTCTCTCAAGGTCAGTTGGTGAATAAT 360
QY 404 ----- 403
Db 361 CGGTTTAAAGGAGTCAAGTATTTTCGCTTGAATACCTTAGCCTCTCTAGGTTATGTGGTT 420
QY 404 ----- 403
Db 421 GTAGTGATAGACAACAGGGGATCTGTGTCCACCGAGGGCTTAAATTTGAAGGCGCTTTAAA 480
QY 404 -----GGTCAAAATGAAATTTGACGATCAGGTGGAAGGACTCCAAATCTAGCTTCT 455
Db 481 TATAAATGGGTCAAATAGAAATTTGACGATCAGGTGGAAGGACTCCAAATCTAGCTTCT 540
QY 456 CGATATGATTTTCACTTAGATCGTGTGGGCATCCACGGCTGGTCTCTATGGAGGATAC 515
Db 541 CGATATGATTTTCACTTAGATCGTGTGGGCATCCACGGCTGGTCTCTATGGAGGATAC 600
QY 516 CTCTCCCTGATGGCATTAATGACAGGTGAGATATCTTTCAGGGTGTCTATTTGCTGGGGCC 575
Db 601 CTCTCCCTGATGGCATTAATGACAGGTGAGATATCTTTCAGGGTGTCTATTTGCTGGGGCC 660
QY 576 CGAGTCACTCTGTGGATCTTATGATACAGGATACAGGACGCTTATATGGTCAACCT 635
Db 661 CGAGTCACTCTGTGGATCTTATGATACAGGATACAGGACGCTTATATGGTCAACCT 720
QY 636 GACGAGAAATGAAACAGGGCTATTACTTAGGATCTGTGGCCATCGAAGACGAAAGTTCCTCC 695
Db 721 GACGAGAAATGAAACAGGGCT-----AAAAAAGGTTTCCCC 756
QY 696 TCTGAACCAAAATCGTTTACTGCTTACATGTTTCTGATGAGATATGTCATTTTGA 755
Db 757 TCTGAACCAAAATCGTTTACTGCTTACATGTTTCTGATGAGATATGTCATTTTGA 816
QY 756 CATACAGTATATATCTAGTGTGTTTATGAGGGCTGGAAGCCATATGATTTACAGATC 815
Db 817 CATACAGTATATATCTAGTGTGTTTATGAGGGCTGGAAGCCATATGATTTACAGATC 876
QY 816 TATCCTCTCAGGAGAGACACAGCATTAAGATTTCTGATCGGAGAACATTTAAGACTGCAT 875
Db 877 TATCCTCTCAGGAGAGACACAGCATTAAGATTTCTGATCGGAGAACATTTAAGACTGCAT 936
QY 876 CTTTGGACCTACCTTCAAGAAACCTTGGATCAGTATTTGCTGCTCTTAAAGTGATATAA 935
Db 937 CTTTGGACCTACCTTCAAGAAACCTTGGATCAGTATTTGCTGCTCTTAAAGTGATATAA 996
QY 936 TTTTGGACCTGTAGAACTCTCTGGTATACACTGCTATTTTAAACCAATGAGAGGTTTA 995
Db 997 TTTTGGACCTGTAGAACTCTCTGGTATACACTGCTATTTTAAACCAATGAGAGGTTTA 1056
QY 996 ATCAACAGAAAAACACAGAAATGATCATCATTTTGTATACCTGCGCATGTAACATCTACTC 1055
Db 1057 ATCAACAGAAAAACACAGAAATGATCATCATTTTGTATACCTGCGCATGTAACATCTACTC 1116
QY 1056 CTGAAATAAATGTGGTGCATGCGGGGCTTACGGTTTGTGGTAGTAATCTAATACCTT 1115
Db 1117 CTGAAATAAATGTGGTGCATGCGGGGCTTACGGTTTGTGGTAGTAATCTAATACCTT 1176
QY 1116 AACCCACATGCTCAAAATCAAATGATATATTTCTGAGAGACCCAGCAATACCAATAG 1175
Db 1177 AACCCACATGCTCAAAATCAAATGATATATTTCTGAGAGACCCAGCAATACCAATAG 1236
QY 1176 AATTACTTAAAAAATAAAAA 1197
Db 1237 AATTACTTAAAAAATAAAAA 1258
```

```
RESULT 3
AF173382
LOCUS Homo sapiens MSTP097 (MST097) mRNA, complete cds. linear HTC 01-AUG-2003
DEFINITION Homo sapiens MSTP097 (MST097) mRNA, complete cds.
ACCESSION AF173382
VERSION AF173382.1 GI:33338001
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE
AUTHORS Xu,H.S., Zhao,B., Tong,Y.K., Sheng,H., Qin,B.M., Liu,Y.Q., Liu,B.,
Wang,X.Y., Zhang,Q., Song,L., Gao,Y., Zhang,C.L., Ye,J., Ji,X.J.,
Liu,B.H., Lu,H., Chen,J.Z., Cai,M.Q., Zheng,W.Y., Teng,C.Y.,
Liu,Q., Yu,L.T., Lin,J., Gong,Q., Zhang,A.M., Gao,R.L. and Hiu,R.T.
Homo sapiens normal aorta mRNA MST097
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 1278)
TITLE Xu,H.S., Zhao,B., Tong,Y.K., Sheng,H., Qin,B.M., Liu,Y.Q., Liu,B.,
Wang,X.Y., Zhang,Q., Song,L., Gao,Y., Zhang,C.L., Ye,J., Ji,X.J.,
Liu,B.H., Lu,H., Chen,J.Z., Cai,M.Q., Zheng,W.Y., Teng,C.Y.,
Liu,Q., Yu,L.T., Lin,J., Gong,Q., Zhang,A.M., Gao,R.L. and Hiu,R.T.
JOURNAL Direct Submission
AUTHORS Submitted (30-JUL-1999) Molecular Medicine Center for
Cardiovascular Disease, Cardiovascular Institute, CAMS & PUMC, 167
Bei Li Shi Lu, Beijing 100037, P.R. China
FEATURES
source Location/Qualifiers
1..1278
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="aorta"
1..1278
/gene="MST097"
591..1010
/gene="MST097"
/codon_start=1
/product="MSTP097"
/protein_id="AAQ13623.1"
/db_xref="GI:33338002"
QY QYVLGSAVMAQAEKFFPFPLHLLHFLDENVHFATSHILLSFLVRAGRPDIQIYP
Db 1 CTGGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCATCTTGTGTCATCAGTCAGC 123
QY 124 ACTGTGACTTCTTTATAAGTATAGTATAGTAACTGACAGAAATCCACACTGTGTGCTCTTT 183
Db 61 ACTGTGACTTCTTTATAAGTATAGTATAGTAACTGACAGAAATCCACACTGTGTGCTCTTT 120
QY 184 ACAAGCTATCAAGTCTGGAAGATGACCAACTTGTGCAAAACAAAGAAATTTTGGGCCCA 243
Db 121 ACAAGCTATCAAGTCTGGAAGATGACCAACTTGTGCAAAACAAAGAAATTTTGGGCCCA 180
QY 244 TTTTGGATTCAGAGGTCTCTCTCTGATATAGTATCTCTCCAGAAATTTCTCTTTGAAA 303
Db 181 TTTTGGATTCAGAGGTCTCTCTCTGATATAGTATCTCTCCAGAAATTTCTCTTTGAAA 240
QY 304 GTACTACTGATTTTACATTTGATGGATGCTCTACAAGCCTCATGATCTACAGCTCGAA 363
Db 241 GTACTACTGATTTTACATTTGATGGATGCTCTACAAGCCTCATGATCTACAGCTCGAA 300
QY 364 AGAAATATCTACTGTGCTGTTTCATATATGTTGGTCTCTCA----- 403
```

Db 301 AGAAATATCTACTGTGCTGTTTCATATATGTTGCTCTCAGGTGCAGTTGGTGAATATC 360
Qy 404 ----- 403
Db 361 GGTTTAAGGAGTCAAGTATTTTCGCTTTGAATACCCCTAGCCTCTCTAGGTATGTGGTTG 420
Qy 404 ----- 403
Db 421 TAGTGATAGACACAGGGGATCTGTGTACCAGGGCTTAAATTTGAAGGCGCTTTAAAT 480
Qy 404 -----GGGTCAAATAGAAATGACGATCAGGTGGAAGACTCCCAATATCTAGCTTCTC 456
Db 481 ATAAATGGGTCAAATAGAAATGACGATCAGGTGGAAGACTCCCAATATCTAGCTTCTC 540
Qy 457 GATATC-ATTTCAATTGACTTATGATCGTGTGGGCATCCACGGCTGGTCTCTATGAGGATAC 515
Db 541 GATATGAATTTTCAATTGACTTATGATCGTGTGGGCATTTCCCGCTGCTCTATGAGATAC 600
Qy 516 CTCTCCCTGATGGCATTAATGCAGAGGTTCAGATATCTTCAGGGTTCCTATGCTGGGCC 575
Db 601 CCTCTCTGATGGCAT--ATGCAGAGGTTCAGATATCTTCAGG--TGCTATGCTGG--C 654
Qy 576 CAGTCACTCTGTGGATCTTCTATGATACAGGATACAGGAAGTTCATATGAGTTCACCT 635
Db 655 CAGTCACTCTGTGATCTCTAT--GAACAGATACAGG-AGTTATATGGGTCACTG 710
Qy 636 GACCAGATTAATACAGGGCTATTACTTAGGATCTCTGGCCATGCAAGCAGAAAGTTCCCC 695
Db 711 ACCAGAAATGAACAGGGCTATTACTTAGGATCTCTGGCCATGCAAGCAGAAAGTTCCCC 770
Qy 696 TCTGAACCAAATCGTTTACTGCTCTTACATGTTTCTCTGGATGAGAAATGTCATTTGCA 755
Db 771 TTTGAACCAAATCGTTTACTGCTCTTACATGTTTCTCTGGATGAGAAATGTCATTTGCA 830
Qy 756 CATACCAATATATTAAGTATTTTCTAGTGGGCTGGAAGCCATATGATTTTACAGATC 815
Db 831 CATACCAATATATTAAGTATTTTCTAGTGGGCTGGAAGCCATATGATTTTACAGATC 890
Qy 816 TATCTCTCAGGAGAGACAGCATAGAGTTCTCTGAATCGGAGAAACATATGAACTGCAT 875
Db 891 TATCTCTCAGGAGAGACAGCATAGAGTTCTCTGAATCGGAGAAACATATGAACTGCAT 950
Qy 876 CTTTTCACATCTCTCAAGAAACCTTGGATCAGTATTTGCTCTCAAAAAGTATATAA 935
Db 951 CTTTTCACATCTCTCAAGAAACCTTGGATCAGTATTTGCTCTCAAAAAGTATATAA 1010
Qy 936 TTTTGACCTGTGAGAACTCTCTGGTATACACTGGCTATTTAACCAATGAGGAGTTTA 995
Db 1011 TTTTGACCTGTGAGAACTCTCTGGTATACACTGGCTATTTAACCAATGAGGAGTTTA 1070
Qy 996 ATCAACAGAAACACAGAAATGATCATCATTCTCTGGTATACCTGCCATGTAACATCTACTC 1055
Db 1071 ATCAACAGAAACACAGAAATGATCATCATTCTCTGGTATACCTGCCATGTAACATCTACTC 1130
Qy 1056 CTGAAATAAATGTGGTGCATGCAAGGGTCTACGGGTTTGTGGTAGTAATCTTAATACCTT 1115
Db 1131 CTGAAATAAATGTGGTGCATGCAAGGGTCTACGGGTTTGTGGTAGTAATCTTAATACCTT 1190
Qy 1116 AACCCACATGCTCAAAATCAAATGATACATATTTCTCTGAGAGCCCAAGCAATACCAATAG 1175
Db 1191 AACCCACATGCTCAAAATCAAATGATACATATTTCTCTGAGAGCCCAAGCAATACCAATAG 1250
Qy 1176 AATTACTAAAAAATAAAAA 1197
Db 1251 AATTACTAAAAAATAAAAA 1272

RESULT 4
CR609512 2292 bp mRNA linear HTC 21-JUL-2004
LOCUS full-length cDNA clone CS0DL005YD02 of B cells (Ramos cell line)
DEFINITION Cot 25-normalized of Homo sapiens (human).
ACCESSION CR609512

CR609512.1 GI:50490319
HTC; CNSLT_CDNA.
Homo sapiens (human)
Homo sapiens
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 2292)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paradise Avenue
2 (bases 1 to 2292)
Genoscope.
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
Location/Qualifiers
1..2292
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DL005YD02"
/issue_type="B cells (Ramos cell line) Cot 25-normalized"
/plasmid="pCMVSPORT_6"

Query Match 65.8%; Score 788; DB 4; Length 2292;
Best Local Similarity 85.5%; Pred. No. 8.6e-207;
Matches 973; Conservative 0; Mismatches 0; Indels 165; Gaps 2;
Qy 1 ATTTTGAAGGCCACCAAGACTCCCTTTAGAGCATCACCTGTAGTACGTTAGCTAA 60
Db 1320 ATTTTGAAGGCCACCAAGACTCCCTTTAGAGCATCACCTGTAGTACGTTAGCTAA 1379
Qy 61 ATCTCTGGAGAGTGAACAAGGCTGACTGACCGTGGCTACTCACATTTCTGTGCATCAGTC 120
Db 1380 ATCTCTGGAGAGTGAACAAGGCTGACTGACCGTGGCTACTCACATTTCTGTGCATCAGTC 1439
Qy 121 AGCACTGTGACTTCTTTTAAAGTAAGTATAGTAACCAAGAAATCCACACTGTGTGTCCTC 180
Db 1440 AGCACTGTGACTTCTTTTAAAGTAAGTATAGTAACCAAGAAATCCACACTGTGTGTCCTC 1499
Qy 181 TTTTACAGCTATCAAGTCTCGAAGATGACCAACTTGCACAAACAAAGAAATTTTGGGCCA 240
Db 1500 TTTTACAGCTATCAAGTCTCGAAGATGACCAACTTGCACAAACAAAGAAATTTTGGGCCA 1559
Qy 241 CCATTTTGGATTCACAGGTCTCTTCTGACTATATCTCTCCAGAAATTTCTCTTTTG 300
Db 1560 CCATTTTGGATTCACAGGTCTCTTCTGACTATATCTCTCCAGAAATTTCTCTTTTG 1619
Qy 301 AAAGTACTACTGGATTACATTTGATGGATGCTCTACAGCCTCATGATCTACAGCCTG 360
Db 1620 AAAGTACTACTGGATTACATTTGATGGATGCTCTACAGCCTCATGATCTACAGCCTG 1679
Qy 361 GAAAGAAATATCTACTGTGCTGTTTCATATATGTTGGTCTCTCAGGGTCAAAATGAAATG 420
Db 1680 GAAAGAAATATCTACTGTGCTGTTTCATATATGTTGGTCTCTCAGGGTCAAAATGAAATG 1722
Qy 421 ACATCAGGTGGAAGACTCCCAATATCTAGCTTCTCGATATGATTTTCATTGACTTAGATC 480
Db 1723 ----- 1722
Qy 481 GTGTGGGCATCCACGGTGGTCTTATGAGGATACCTCTCTCCCTGATGGCATTATGATGAGA 540

Db 1723 ----- 1722

Qy 541 GGTACAGATATCTTCAGGGTTGCTATTGCTGGGGCCAGTCACCTCTCTGGATCTTCTATG 600

Db 1723 -----GGTTGCTATTGCTGGGGCCAGTCACCTCTCTGGATCTTCTATG 1766

Qy 601 ATACAGGATACACGGAACGTTATATGGGTACCTTGACCCAGAGATGAACAGGGCTATTACT 660

Db 1767 ATACAGGATACACGGAACGTTATATGGGTACCTTGACCCAGAGATGAACAGGGCTATTACT 1826

Qy 661 TAGGATCTGGCCATGCAAGCAGAAAAGTTCCCTCTGACCAAAATCGTTTACTGCTCT 720

Db 1827 TAGGATCTGGCCATGCAAGCAGAAAAGTTCCCTCTGACCAAAATCGTTTACTGCTCT 1886

Qy 721 TACATGGTTTCCTGGATGAGAAATGTCATTTTGGACATACACAGTATATTACTGAGTTTTT 780

Db 1887 TACATGGTTTCCTGGATGAGAAATGTCATTTTGGACATACACAGTATATTACTGAGTTTTT 1946

Qy 781 TAGTGAGGCTGGAAGCCATATGATTTACAGATCTATCTCAGGAGAGACACAGCATAA 840

Db 1947 TAGTGAGGCTGGAAGCCATATGATTTA-----CAGGAGAGACACAGCATAA 1994

Qy 841 GAGTTCTGANTCGGAGAACATTATGACTGCATCTTTTGCACCTACTCTTCAAGAAACC 900

Db 1995 GAGTTCTGANTCGGAGAACATTATGACTGCATCTTTTGCACCTACTCTTCAAGAAACC 2054

Qy 901 TTGGATCACGATTTGCTGCTCTAAAGTGATATATTTTGACCTGTGTAGAACTCTCTGG 960

Db 2055 TTGGATCACGATTTGCTGCTCTAAAGTGATATATTTTGACCTGTGTAGAACTCTCTGG 2114

Qy 961 TATACACTGGCTATTTAACCAATGAGGAGTTTAAATCAACAGAAAACACAGAAATGATC 1020

Db 2115 TATACACTGGCTATTTAACCAATGAGGAGTTTAAATCAACAGAAAACACAGAAATGATC 2174

Qy 1021 ATCACATTTTGATACCTGCATGTAAACATCTACTCTGAAAATAAATGTGGTGCATGCA 1080

Db 2175 ATCACATTTTGATACCTGCATGTAAACATCTACTCTGAAAATAAATGTGGTGCATGCA 2234

Qy 1081 GGGGCTACGGTTTGGGTAGTAATCTAATACCTTAAACCCACATGCTCAAAATCAA 1138

Db 2235 GGGGCTACGGTTTGGGTAGTAATCTAATACCTTAAACCCACATGCTCAAAATCAA 2292

RESULT 5

AL542617

LOCUS

DEFINITION AL542617 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE014YN22

5-PRIME, mRNA sequence.

AL542617

AL542617.3 GI:45718193

EST.

Homd sapiens (human)

Homd sapiens

Homd sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.

1 (bases 1 to 864)

Li,W.B., Gruber,C., Jessee,J. and Polayes,D.

Full-length cDNA libraries and normalization

Unpublished (2001)

On Feb 15, 2001 this sequence version replaced gi:30547946.

Contact: Genoscope

Genoscope - Centre National de Sequencage

2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized. Library was constructed by Life Technologies, a division of Invitrogen.

This sequence belongs to sequence cluster 7542.r

For more information about this cluster, see

http://www.genoscope.cns.fr/cdna?s=CS0DE014DG11QPI&c=7542.r.

FEATURES

source

Location/Qualifiers

1. 864

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CS0DE014YN22"

/tissue_type="PLACENTA"

/clone_lib="Homo sapiens PLACENTA"

/notes="Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

ORIGIN

Query Match 64.2%; Score 768; DB 1; Length 864;

Best Local Similarity 98.9%; Pred. No. 2.3e-201;

Matches 787; Conservative 6; Mismatches 1; Indels 2; Gaps 2;

Qy 404 GGGTCAAAATAGAAATTCAGATCAGGTGGAAGACTCCAAATATCTAGCTTCTCGATATGA 463

Db 64 GGGTCAAAATAGAAATTCAGATCAGGTGGAAGACTCCAAATATCTAGCTTCTCGATATGA 123

Qy 464 TTTTCAATTGACTTAGATCGTGTGGGCATCCACGGCTGCTCTATGAGGATACCTCTCCCT 523

Db 124 TTTTCAATTGACTTAGATCGTGTGGGCATCCACGGCTGCTCTATGAGGATACCTCTCCCT 183

Qy 524 GATGGCATTAAATGCAGAGGTTCAGATATCTTCAGGGTTCCTATGCTGGGGCCCCAGTCCAC 583

Db 184 GATGGCATTAAATGCAGAGGTTCAGATATCTTCAGGGTTCCTATGCTGGGGCCCCAGTCCAC 243

Qy 584 TCTGTGGATCTTCTATGATACAGGATACACGGAACTTATATGCTACCTGACCCAGAA 643

Db 244 TCTGTGGATCTTCTATGATACAGGATACACGGAACTTATATGCTACCTGACCCAGAA 303

Qy 644 TGAACAGGGCTATTACTTAGGATCTGTGGCATGCAAGCAGAAAAGTTCCCTCTGACCC 703

Db 304 TGAACAGGGCTATTACTTAGGATCTGTGGCATGCAAGCAGAAAAGTTCCCTCTGACCC 363

Qy 704 AAATCGTTTACTGCTCTTACATGTTTCTGATGAGAAATGCCATTTTGCACATACCAG 763

Db 364 AAATCGTTTACTGCTCTTACATGTTTCTGATGAGAAATGCCATTTTGCACATACCAG 423

Qy 764 TATATTACTGAGTTTTTTAGTGAGGGCTGAAAGCCATATGATTTACAGATCTATCTCTCA 823

Db 424 TATATTACTGAGTTTTTTAGTGAGGGCTGAAAGCCATATGATTTACAGATCTATCTCTCA 483

Qy 824 GGAGAGACACAGCATTAAGAGTTCTGAAATCGGGAGAAATTAATGAACTGCACTTTTGA 883

Db 484 GGAGAGACACAGCATTAAGAGTTCTGAAATCGGGAGAAATTAATGAACTGCACTTTTGA 543

Qy 884 CTACCTTCAAGAAAACCTTGGATCAGTATTTGCTGCTCTTAAAGTGATATAATTTTGACC 943

Db 544 CTACCTTCAAGAAAACCTTGGATCAGTATTTGCTGCTCTTAAAGTGATATAATTTTGACC 603

Qy 944 TGTGTAGAACTCTCTGGTATACACTGGCTATTTAAACCAATGAGGAGTTTAAATCAACAG 1003

Db 604 TGTGTAGAACTCTCTGGTATACACTGGCTATTTAAACCAATGAGGAGTTTAAATCAACAG 663

Qy 1004 AAAACACAGAAATTCATCATCAATTTTATCTGCTGATGAACTTACTCTCTGAAAAAT 1063

Db 664 AAAACACAGAAATTCATCATCAATTTTATCTGCTGATGAACTTACTCTCTGAAAAAT 723

Qy 1064 AAATGTGGTCCCATGAGGGCTCTAGCGTTTGGTGTAGTAAATCTAATACCTTAAACCCAC 1123

Db 724 AAATGTGGTCCCATGAGGGGKCTACGGTTTGGTGTAGTAAATCTAATACCTTAAACCCAC 783

Qy 1124 ATGCTCAAAATCAAA-TGATACATATTCC-TGAGAGACCCAGCAATACCAAGAAATTAC 1181

Db 784 ATGCTCAAAATCAAAATTGATACATATTCTTTGAGAGCCAGCAATACCAAGAAATTAC 843

Qy 1182 TAAAAAATCAAAATCAAAATCAAAATCAAAATCAAAATCAAAATCAAAATCAAAATCAAA 1197

|||||

```

Db      844 CAAAAAAAAAAAAA 859

RESULT 6
BU687687/c
LOCUS
DEFINITION
UT-CF-EC1-adu-o-02-0-UI_s1 UI-CF-EC1 Homo sapiens cDNA clone
UT-CF-EC1-adu-o-02-0-UI 3', mRNA sequence.
ACCESSION
BU687687
VERSION
BU687687.1 GI:23543781
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
1 (bases 1 to 804)
Bonaldo,M.F., Lennon,G. and Soares,M.B.
Normalisation and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
8889548
Contact: McCray, PB
McCray Lab
University of Iowa
2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
Tel: 319 356 4866
Fax: 319 356 7171
Email: paul-mccray@uiowa.edu
Tissue Procurement: Dr. M. J. Welsh, University of Iowa
CDNA Library preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.regen.com) or from Open Biosystems
(www.openbiosystems.com).
Seq primer: M13 FORWARD
POLYA=yes.

FEATURES
    source
        1..804
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="UI-CF-EC1-adu-o-02-0-UI"
            /tissue_type="Lung"
            /dev_stages="Adult and Fetal"
            /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
            /clone_libs="UI-CF-EC1"
            /note="Organ: Lung; Vector: p7T3-Pac (Pharmacia) with a
            modified polylinker; Site 1: EcoR I; Site 2: Not I;
            UI-CF-EC1 is a normalized cDNA library containing the
            following tissue(s): Normal lung from adult and from fetal
            day 64, day 87, week 19 and week 42. The library was
            constructed according to Bonaldo, Lennon and Soares,
            Genome Research, 6:791-806 1996, First strand cDNA,
            synthesis was primed with an oligo-dT primer containing a
            Not I site. Double stranded cDNA was ligated to an EcoR I
            adaptor digested with Not I, and cloned directionally
            into p7T3-Pac vector. The oligonucleotide used to prime
            the synthesis of first-strand cDNA contains a library tag
            sequence that is located between the Not I site and the
            GATC tail. The sequence tag for this library is
            AAGTCTTAC.
            TAG-TISSUE=Normal Lung Epithelial Cells Tissue nos 369-371
            and 380-383
            TAG-LIB=UI-CF-EC1
            TAG_SEQ=AAGTCTTAC"

ORIGIN
Query Match      64.0%; Score 766.2; DB 5; Length 804;
Best Local Similarity 99.4%; Pred. No. 7.2e-201;
Matches 790; Conservative 0; Mismatches 3; Indels 2; Gaps 2;

1183 AAAAAAAAAAAAAA 1197
18 AAAAAAAAAAAAAA 4

RESULT 7
AY411615
LOCUS
DEFINITION
Homo sapiens DPP8 gene, VIRTUAL TRANSCRIPT, partial sequence,
Genomic survey sequence.
AY411615
ACCESSION
AY411615.1 GI:39767583
VERSION
KEYWORDS
GSS.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
1 (bases 1 to 2649)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,

403 AGGTCGAATAGAAATTGACGATCAGGTGAGAGGCTCAATATCTAGCTTCTCGATATG 462
796 ATGGTCAATAGAAA-TGACGATCA-GTGAAGGACTCCAATATCTAGCTTCTCGATATG 739
463 ATTTCAATGACTTAGATTCGTGTGGGATCCACGGCTGGTCTTATGGAGGATACCTCTCCC 522
738 ATTTCAATGACTTAGATTCGTGTGGGATCCACGGCTGGTCTTATGGAGGATACCTCTCCC 679
523 TGATGGCATTAAATGCAGAGGTGAGATATCTTCAGGGTTGCTATTGCTGGGGCCCCAGTCA 582
678 TGATGGCATTAAATGCAGAGGTGAGATATCTTCAGGGTTGCTATTGCTGGGGCCCCAGTCA 619
583 CTCTGTGGATCTTCTATGATACAGGATACACGGAACTTATATGGGTCACTCTGACAGA 642
618 CTCTGTGGATCTTCTATGATACAGGATACACGGAACTTATATGGGTCACTCTGACAGA 559
643 ATGAACAGGGCTATTACTTAGGATCTGTGGCCATCAAGAGCAAAAGTTCCCTCTGAAC 702
558 ATGAACAGGGCTATTACTTAGGATCTGTGGCCATCAAGAGCAAAAGTTCCCTCTGAAC 499
703 CAATCGTTTACTGCTCTTACATGTTTCTCGATGAGAATGTCATTTGCACATACCA 762
498 CAATCGTTTACTGCTCTTACATGTTTCTCGATGAGAATGTCATTTGCACATACCA 439
763 GTATATTACTGAGTTTCTTAGTGGGCTGGAAAGCCATATGATTACAGATCTATCTC 822
438 GTATATTACTGAGTTTCTTAGTGGGCTGGAAAGCCATATGATTACAGATCTATCTC 379
823 AGGAGACACACAGCATACAGTTCTCTGAATCGGAGAACATATGAATGATCTGCTTTC 882
378 AGGAGACACACAGCATACAGTTCTCTGAATCGGAGAACATATGAATGATCTGCTTTC 319
883 ACTACCTTCAGAAAAACCTTGGATCAGGTATGCTGCTCTAAAAGTGATATAATTTTGAC 942
318 ACTACCTTCAGAAAAACCTTGGATCAGGTATGCTGCTCTAAAAGTGATATAATTTTGAC 259
943 CTGTGTAGAACTCTCTGGTATACACTGGCTATTAAACCAATAGAGGAGTTAATCAACA 1002
258 CTGTGTAGAACTCTCTGGTATACACTGGCTATTAAACCAATAGAGGAGTTAATCAACA 199
1003 GAAAAACACAGAAATTGATCATCATCTTTTGATACCTGCCATGTAAACATCTACTCTGAAA 1062
198 GAAAAACACAGAAATTGATCATCATCTTTTGATACCTGCCATGTAAACATCTACTCTGAAA 139
1063 TAAATGTGTGGTCATGCGGGGTCTACGGTTTGTGGTAGTAACTCTAATACCTTAACCCCA 1122
138 TAAATGTGTGGTCATGCGGGGTCTACGGTTTGTGGTAGTAACTCTAATACCTTAACCCCA 79
1123 CATGCTCAAAATCAATGATACATATTCCTGAGAGCCAGCAATACCAATAGAAATTAAT 1182
78 CATGCTCAAAATCAATGATACATATTCCTGAGAGCCAGCAATACCAATAGAAATTAAT 19
1183 AAAAAAAAAAAAAA 1197
18 AAAAAAAAAAAAAA 4

```

Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D. and Cargill, M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
Science 302 (5652), 1960-1963 (2003)
14671302
2 (bases 1 to 2649)
Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D. and Cargill, M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source
Location/Qualifiers
1..2649
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..>2649
/gene="DPP8"
/locus_tag="HCM4283"
ORIGIN
Query Match 63.4%; Score 758.4; DB 10; Length 2649;
Best Local Similarity 84.7%; Pred. No. 1.4e-198;
Matches 916; Conservative 0; Mismatches 19; Indels 147; Gaps 1;
QY 1 ATTTTGAAGGACCAAAAGACTCCCTTTAGAGCATCACCTGTACGTAGTACGTAA 60
DB 1568 ATTTTGAAGGACCAAAAGACTCCCTTTAGAGCATCACCTGTACGTAGTACGTAA 1627
QY 61 ATCTGGAGAGTGACAAGGCTGACTGACCGTGGCTACTCACAATCTTCTGTCATCAGTC 120
DB 1628 ATCTGGAGAGTGACAAGGCTGACTGACCGTGGCTACTCACAATCTTCTGTCATCAGTC 1687
QY 121 AGCACTGTGACTCTTTTATAGTAAGTATAGTAACAGAGAAATCCACACTGTGTGTC 180
DB 1688 AGCACTGTGACTCTTTTATAGTAAGTATAGTAACAGAGAAATCCACACTGTGTGTC 1747
QY 181 TTACAAGCTATCAAGTCTGAAGATGACCCAACTTGCAGAAACAAAGAAATTTGGGCCA 240
DB 1748 TTACAAGCTATCAAGTCTGAAGATGACCCAACTTGCAGAAACAAAGAAATTTGGGCCA 1807
QY 241 CCAATTTGGATTCAGAGGCTCTCTTCTGACTATCTCTCCAGAGAAATTTCTCTTTG 300
DB 1808 CCAATTTGGATTCAGAGGCTCTCTTCTGACTATCTCTCCAGAGAAATTTCTCTTTG 1867
QY 301 AAAGTACTACTGGATTTACATTTGATGGGATGCTCTACAAGCTCATGATCTACAGCTG 360
DB 1868 AAAGTACTACTGGATTTACATTTGATGGGATGCTCTACAAGCTCATGATCTACAGCTG 1927
QY 361 GAAAGAAATATCTACTGTGCTGTTCATATATATGCTGCTCTCA----- 403
DB 1928 GAAAGAAATATCTACTGTGCTGTTCATATATATGCTGCTCTCAAGGATGTTGGTAATA 1987
QY 404 ----- 403
DB 1988 ATCGGTTTAAAGGAGTCAAGATATTTCCGCTTGAATACCCTAGCCTCTTAGTGTATGTGG 2047
QY 404 ----- 403
DB 2048 TTGTAGTATAGACAAACAGGGATCTGTCCACCGAGGCTTAAATTTGAAGCGGCTTTTA 2107
QY 404 -----GGGTCAAAATGAAATTCACGATCAGGTGGAAGACTTCCAATATCTAGCTT 453
DB 2108 AATATAAAATGGGTCAATAGAAATTCACGATCAGGTGGAAGACTTCCAATATCTAGCTT 2167
QY 454 CTCGATATGATTTTCATTGACTTATAGATCGTGTGGGCATCCACGGCTGGTCTTATGAGGAT 513
DB 2168 CTCGATATGATTTTCATTGACTTATAGATCGTGTGGGCATCCACGGCTGGTCTTATGAGGAT 2227

QY 514 ACCTCTCCCTGATGGCATTAAATCAGAGTTCAGATATCTTCAGGTTGCTATTGCTGGG 573
DB 2228 ACCTCTCCCTGATGGCATTAAATCAGAGTTCAGATATCTTCAGGTTGCTATTGCTGGG 2287
QY 574 CCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAAGCTTATATGGTCA 633
DB 2288 CCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAAGCTTATATGGTCA 2347
QY 634 CTGACCAAGATGAACACGGCTATTACTTAGGATCTCTGGCCATGCAAGCAGAAAAATTC 693
DB 2348 CTGACCAAGATGAACACGGCTATTACTTAGGATCTCTGGCCATGCAAGCAGAAAAATTC 2407
QY 694 CCTCTGAACCAAAATCGTTTACTCTCTTACATGGTTTCTCTGATGAGAAATGTCATTTG 753
DB 2408 NNNNNNACCAAAATCGTTTACTCTCTTACATGGTTTCTCTGATGAGAAATGTCATTTG 2467
QY 754 CACATACCAAGTATATTAATCTAGTCTTCTTCTAGTGGCTGGAAGCCATATGTTTACAGA 813
DB 2468 CACATACCAAGTATATTAATCTAGTCTTCTTCTAGTGGCTGGAAGCCATATGTTTACAGA 2527
QY 814 TCTATCTCTCAGGAGACACAGCATTAAGAGTTCTCTGAATCGGAGAACATATATGAATGC 873
DB 2528 TCTATCTCTCAGGAGACACAGCATTAAGAGTTCTCTGAATCTCGAGAACATATATGAATGC 2587
QY 874 ATCTTTTGCACCTACCTTCAAGAAAAACCTTGGATCAGTATTTCTCTTAAAGTGATAT 933
DB 2588 ATCTTTTGCACCTACCTTCAAGAAAAACCTTGGATCAGTATTTCTCTTAAAGTGATAT 2647
QY 934 AA 935
DB 2648 AA 2649
RESULT 8
AI917735/c
LOCUS
DEFINITION
t11d07.x1 NCI CGAP GC6 Homo sapiens cDNA clone IMAGE:2240461 3'
similar to TR:075273 075273 R26984_1 ; mRNA sequence.
ACCESSION AI917735
VERSION AI917735.1 GI:5637590
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 760)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 1041 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 444.
location/Qualifiers
source
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clones="IMAGE:2240461"
/tissue_type="pooled germ cell tumors"


```
Db 1988 ATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCCCTAGCCTCTCTAGGTTATGTGG 2047
Qy 404 ----- 403
Db 2048 TTGTAGTGATAGACAACAGGGGATCTGTACCGAGGGCTTAAATTTTGAAGGCGCCTTTA 2107
Qy 404 -----GGGTCAAATAGAAATTTGACGATCAGGTGGAAGGACTCCAATATCTAGCTT 453
Db 2108 AATATAAATGGGTCAATAGAAATTTGACATCAGTGGAGGACTCCAATATCTAGCTT 2167
Qy 454 CTCGATATGATTTTCATGACTTAGATCGGTGGGCATCCACGGCTGGTCTCTATGGAGGAT 513
Db 2168 CTCGATATGATTTTCATGACTTAGATCGGTGGGCATCCACGGCTGGTCTCTATGGAGGAT 2227
Qy 514 ACCTCTCCCTGATGGCATTAATGCAGAGGTGAGATATCTTCAGGGTTGCTATTGCTGGGG 573
Db 2228 ACCTCTCCCTGATGGCATTAATGCAGAGGTGAGATATCTTCAGGGTTGCTATTGCTGGGG 2287
Qy 574 CCCAGTCTCTGTGGATCTTCTATGATACAGGATACACGGAACGTTATATGGTCCACC 633
Db 2288 CCCAGTCTCTGTGGATCTTCTATGATACAGGATACACGGAACGTTATATGGTCCACC 2347
Qy 634 CTGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCCATCGAAGCAGAAAAGTTCC 693
Db 2348 CTGACCAAGATGAACANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2407
Qy 694 CCTCTGAACCAATCGTTTACTGCTCTTACATGTTTCTCTGGATGAGAATGTCCATTTTG 753
Db 2408 NNNNNNACCAATCGTTTACTGCTCTTACATGTTTCTCTGGATGAGAATGTCCATTTTG 2467
Qy 754 CACATACCAAGTATTTACTGAGTGTGTTTTAGTGAGGGCTGGAAAGCCATGATTACAGA 813
Db 2468 CACATACCAAGTATTTACTGAGTGTGTTTTAGTGAGGGCTGGAAAGCCATGATTACAGA 2527
Qy 814 TCTATCTCGAGAGACACAGCATGAAGTTCCTGAATCGGAGACATATGAAGTGC 873
Db 2528 TCTATCTCGAGAGACACAGCATGAAGTTCCTGAATCTGGAGAACATTTATGAAGTGC 2587
Qy 874 ATCTTTTGCACTACCTTCAAGAAACCTTGGATCAGTATGCTGCTCTAAAGTGATAT 933
Db 2588 ATCTTTTGCACTACCTTCAAGAAACCTTGGATCAGTATGCTGCTCTAAAGTGATAT 2647
Qy 934 AA 935
Db 2648 AA 2649

RESULT 10
AK029788
LOCUS AK029788 5517 bp mRNA linear HTC 03-APR-2004
DEFINITION Mus musculus adult male testis cDNA, RIKEN full-length enriched
library, clone:4930560C15 product:DIPEPTIDYL PEPTIDASE 8 homolog
(Homo sapiens), full insert sequence.
ACCESSION AK029788
VERSION AK029788.1 GI:26081520
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
1
REFERENCE 1
AUTHORS Carninci, P. and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
PUBMED 10349636
2
REFERENCE 2
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
PUBMED 11042159
```

REFERENCE
AUTHORS

3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Sakumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
11076861

JOURNAL
PUBMED

4 The RIKEN Genome Exploration Research Group Phase II Team and the
PANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)

REFERENCE
AUTHORS

5 The PANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 5517)

JOURNAL
PUBMED

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoaka, T., Hirozane, T.,
Hori, P., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, K., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp,
URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)

COMMENT

cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.

Please visit our web site for further details.

URL: http://genome.gsc.riken.jp/

URL: http://fantom.gsc.riken.jp/.

FEATURES

source

1..5517
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:4930560C15"
/db_xref="taxon:10090"
/clone="4930560C15"
/sex="male"
/tissue_type="testis"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
2221..4837
/note="DIPEPTIDYL PEPTIDASE 8 homolog [Homo sapiens]
(SPTLQ9HBM5, evidence: FASTY, 95.7%ID, 100*length,
match=2649)
..putative"

misc_feature

ORIGIN

Query Match 58.9%; Score 704.8; DB 4; Length 5517;
Best Local Similarity 76.1%; Pred. No. 1.3e-183;

Matches 1008; Conservative 0; Mismatches 157; Indels 159; Gaps 5;		
QY	1	ATTTTGAAGCACAAGACTCCCTTTAGACATCACCTGTAGTGTAGTACGTAA 60
Db	3756	ACTTTGAAGCACAAGACTCTCTTTGAGCATCACCTGTAGTGTAGTGTATGCAA 3815
QY	61	ATCTGTGAGGTGACAGGCTGACGACGTTGCTACTCACTTTCTGTGTGACATGTC 120
Db	3816	ACCTGTGAGGTGAGGCTGACGACGTTGCTACTCACTTTCTGTGTGCTGCTCAGCC 3875
QY	121	AGCACTGTGACTCTTTTAAAGTAAGTATAGTAACCAAGAAATCCACATGTGTGCC 180
Db	3876	GGCAATTGTGACTCTTTTAAAGTAAGTATAGTAACCAAGAAATCCACATGTGTGCC 3935
QY	181	TTTCAAGCTATCAAGTCTGAGATGACCCAACTTGCAGAAACAAAGGAAATTTTGGGCCA 240
Db	3936	TCTACAACCTCTCAAGTCTGAGATGACCCAACTTGCAGAAACAAAGGAAATTTTGGGCCA 3995
QY	241	CCATTTTGGATTTCAGCAGGCTCTCTTCTGACTATCTCTCCAGAAATTTTCTTTTG 300
Db	3996	CCATTTTGGATTTCAGCAGGCTCTCTTCTGACTATCTCTCCAGAAATTTTCTTTTG 4055
QY	301	AAAGTACTCTGATTTTCAATTTGATGGATGCTCTCAAGCCTCATGATCTACAGCTG 360
Db	4056	AAAGTACTCTGATTTTCAATTTGATGGATGCTCTCAAGCCTCATGATCTACAGCTG 4115
QY	361	GAAGAAATATCTACTGTGCTGTTCATATATGTTGGTCCCTCA----- 403
Db	4116	GAAGAAATATCTACTGTGCTGTTCATATATGTTGGTCCCTCAAGGTCAGCTGGTGAACA 4175
QY	404	----- 403
Db	4176	ATCGGTTTAAAGGTCAGTATTTCCGCTGAACACCTCGCTCTCTGGTTATGTGG 4235
QY	404	----- 403
Db	4236	TTGTGTGTATAGACACAGGGATCTGTCCAGGACTTAAATTTGAAGGCGCCTTTA 4295
QY	404	-----GGTCAATAGAAATGACATCAGGTGGAAGGACTCCCAATATCTAGCTT 453
Db	4296	AATATAAATGGGTCAATAGAAATCGATGATCAAGTGAAGGACTCCAGTACTAGCAT 4355
QY	454	CTCGATATGATTTCTAGTCTAGTGTGTGGGATCCAGCGTGTCTCTATGGAGGAT 513
Db	4356	CTCAGTATGACTTTCTAGTCTAGTGTGTGGGATCCAGCGTGTCTCTATGGTGGCT 4415
QY	514	ACCTCTCCCTGATGGCATTAATGCAGAGGTGAGATATCTTCAGGGTTGCTATTGCTGGGG 573
Db	4416	ACCTCTCCCTGATGGCATTAATGCAGAGGTGAGATATCTTCAGGGTTGCTATTGCTGGGG 4475
QY	574	CCCAGTCTACTGTGGATCTTCTATGATACAGGATACACGGAACGTTATATGGGTCAAC 633
Db	4476	CCCAGTCTACTGTGGATCTTCTATGATACAGGATACACGGAACGTTATATGGGTCAAC 4535
QY	634	CTGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAGTTCC 693
Db	4536	CTGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAGTTCC 4595
QY	694	CCTCTGAACCAATCGTTTACTGCTTCTACATGTTTCTCGATGAGAATGTCCATTTTG 753
Db	4596	CCTCAGAACCAACCGTTTACTCTCTTGCATGGTTCTTGGATGAGAATGTTCACCTTG 4655
QY	754	CACATACCATGATATTACTGAGTTTTTTAGTGAAGGCTGGAAAGCCATATGATTACAGA 813
Db	4656	CACACACCATGATATTACTGAGTTTTTTAGTGAAGGCTGGAAAGCCATATGATTACAGA 4715
QY	814	TCTATCTCAGGAGACACATAGAGTTCTGATCGGAGACATTTATGACTGC 873
Db	4716	TCTATCTCAGGAGACACATAGAGTTCTGATCGGAGACATTTATGACTGC 4775
QY	874	ATCTTTTGTACCTTCAAGAAACCTTGGATCACCTATGCTGCTCAAGATCATAT 933
Db	4776	ACCTGTCTCACTACCTTCAGGAAACCTTGGATCCGATTCGCTCTGAAAGTATAT 4835

QY	934	AAATTTGACCTGTGTAGAACTCTCTGTGTATACACTGGGCTATTTAAACCAATGAGAGGTT 993
Db	4836	AACTCTGACCATGTAGAACT---TGGTGTACACTG---TCTAACCAACAAAGAGGGT 4888
QY	994	TAATCAACAGAAACACAGAAATGATCATCATTTTGTATCTACCTGCTGCAATCTTAC 1053
Db	4889	TAGCAATAGAGAGGACAGAGTTGATGAACCTTACTTTTGTACCTGCTGCAATCTTAT 4948
QY	1054	TCCTGAAAATTAATTTGGTGGCCATGACGGGTCTACGGTTTGTGG-TAGTAACTTAATAC 1112
Db	4949	TTCTGAAAGTAAATTTGGTGGCCATGACAGAAATCTGCAGTTTCATGATAGTAGTACTCTAATAC 5008
QY	1113	CTTAACCCCATGCTCAAAATCAAAATGATACATATCTCTGAGAGCCACCAATACCAT 1172
Db	5009	CTTAAC---CATGCACATAAACTGAGTGACATGTTCCCAAGGATTGAACAGTACCTG 5064
QY	1173	AAGA 1176
Db	5065	AGGA 5068
RESULT 11		
AK016546		
LOCUS	AK016546	3143 bp mRNA linear HTC 03-APR-2004
DEFINITION	Mus musculus adult male testis cDNA, RIKEN full-length enriched library, clone:4932434F09 product:DIPEPTIDYL PEPTIDASE 8 homolog [Homo sapiens], full insert sequence.	
ACCESSION	AK016546	1 GI:12855334
VERSION	AK016546	1
KEYWORDS	HTC; CAP trapper.	
SOURCE	Mus musculus (house mouse)	
ORGANISM	Mus musculus	
REFERENCE	Carninci,P. and Hayashizaki,Y. High-efficiency full-length cDNA cloning Meth. Enzymol. 303, 19-44 (1999) 10349636	
AUTHORS	Carninci,P. and Hayashizaki,Y.	
TITLE	High-efficiency full-length cDNA cloning	
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)	
PUBMED	10349636	
REFERENCE		
AUTHORS	Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.	
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes	
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)	
PUBMED	11042159	
REFERENCE		
AUTHORS	Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsuunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kita,A. and Hayashizaki,Y.	
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer	
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)	
PUBMED	11076861	
REFERENCE		
AUTHORS	The RIKEN Genome Exploration Research Group Phase II Team and the PANTOM Consortium	
TITLE	Functional annotation of a full-length mouse cDNA collection	
JOURNAL	Nature 409, 685-690 (2001)	
REFERENCE		
AUTHORS	The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase II Team	
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs	
JOURNAL	Nature 420, 563-573 (2002)	
REFERENCE	6 (bases 1 to 3143)	
AUTHORS	Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Arai,A., Aono,H.,	

/tissue_type="choriocarcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NH_MGC_21"
/note="Organ: placenta; Vector: pOTB7; Site 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."

ORIGIN

Query Match 56.3%; Score 674.2; DB 2; Length 910;
Best Local Similarity 94.7%; Pred. No. 2.4e-175;
Matches 730; Conservative 0; Mismatches 38; Indels 3; Gaps 3;

QY 404 GGGTCAAAATAGAAATGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGA 463
DB 139 GGGTCAAAATAGAAATGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGA 198
QY 464 TTTCATTGACTTAGATCGTGTGGGCATCCACGGCTGGTCTTATGGAGGATACCTCTCCCT 523
DB 199 TTTCATTGACTTAGATCGTGTGGGCATCCACGGCTGGTCTTATGGAGGATACCTCTCCCT 258
QY 524 GATGGCATTAAATGCAGAGGTGATATCTTCAGGGTTCCTATTCGTGGGGCCCCAGTCA 583
DB 259 GATGGCATTAAATGCAGAGGTGATATCTTCAGGGTTCCTATTCGTGGGGCCCCAGTCA 318
QY 584 TCTGTGGATCTTCTATGATACAGGATACACGGACGTTATATGGGTGACCTGACCCAGAGAA 643
DB 319 TCTGTGGATCTTCTATGATACAGGATACACGGACGTTATATGGGTGACCTGACCCAGAA 378
QY 644 TGAACAGGGCTATTACTAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCTCTGAACC 703
DB 379 TGAACAGGGCTATTACTAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCTCTGAACC 438
QY 704 AATCGTTTACTGCTTTACATGGTTTCTCGATGAGAGTCTCCATTTTGCATATCTATCTCA 823
DB 439 AATCGTTTACTGCTTTACATGGTTTCTCGATGAGAGTCTCCATTTTGCATATCTATCTCA 498
QY 764 TATATTACTAGTGTTTTATGTCAGGCTGGAAGCCATATGATTTACAGATCTATCTCA 823
DB 499 TATATTACTAGTGTTTTATGTCAGGCTGGAAGCCATATGATTTACAGATCTATCTCA 558
QY 824 GGAGACACAGCATAGAAGTTCCTGAATCGGAGAACATTTATGAATGCAATCTTTTGA 883
DB 559 GGAGACACAGCATAGAAGTTCCTGAATCGGAGAACATTTATGAATGCAATCTTTTGA 618
QY 884 CTACCTTCAAGAAAACCTTGGATCAGTATTCGTCTCTAAAGTGATATAATTTGACC 943
DB 619 CTACCTTCAAGAAAACCTTGGATCAGTATTCGTCTCTAAAGTGATATAATTTGACC 678
QY 944 TGTGTAGACTCTCTGGTATACATCGGCTATTAAACAAATGAGAGGTTTAAATCAACAG 1003
DB 679 TGTGTAGACTCTCTGGTATACATCGGCTATTAAACAAATGAGAGGTTTAAATCAACAG 738
QY 1004 AAAACACAGAAATGATCATCATATTTGATATCTCCATGTAAATCTACTCTCTGAAAT 1063
DB 739 GAAACACAGAAATGATCATCATATTTGATATCTCTGAGGAGTGAAGTAATCAATACCTTGA 797
QY 1064 AA-ATGTGTGCCATGCAGGGGTCTACGGTTTGTGGTAGTAAATCTAATACCTTAACCCCA 1122
DB 798 AATGTGGTGGCCATGCAGGGGTCTACGGGAGTGAAGTAATCAATACCTTAAACCCGA 857
QY 1123 CATGCTCAAAATCAAAATGATACATATTTCTTGAGAGACCCAGCAATACCATTA 1173
DB 858 AATGGTTCAATCAATTGGT-CATTTTCTGAGAAAACAGCAGTAAACCATTA 907

RESULT 14
BU622228/c
LOCUS BU622228 708 bp mRNA linear EST 23-SEP-2002

DEFINITION UI-H-FHI-bfu-d-21-0-UI.s1 NCI CGAP FHI Homo sapiens cDNA clone
UI-H-FHI-bfu-d-21-0-UI 3', mRNA sequence.
ACCESSION BU622228 GI:23288443
VERSION BU622228.1
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 708)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITILE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: James Martin
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, bento-soares@uiowa.edu
Seq primer: M13 FORWARD
POLYA-Yes.

Location/Qualifiers
1..708
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-H-FHI-bfu-d-21-0-UI"
/tissue_type="Cell Line"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI CGAP FHI"
/notes="Organ: Chondrosarcoma; Vector: p7T3-Pac
(Pharmacia) with a modified polylinker; Site 1: EcoR I;
Site 2: Not I; NCI CGAP FHI is a normalized cDNA library
obtained from a cell line derived from grade I
chondrosarcoma tissue. The library was constructed and
normalized according to Bonaldo, Lennon and Soares, Genome
Research, 6:791-806, 1996. First strand cDNA synthesis was
primed with an oligo-dT primer containing a Not I site.
Double stranded cDNA was ligated to an EcoR I adaptor,
digested with Not I, and cloned directionally into
p7T3-Pac vector. The oligonucleotide used to prime the
synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(dT)18 tail. The sequence tag for this library is
AGAATCCGGC. The cell line was provided by Dr. James Martin
from the University of Iowa
TAG TISSUE=Human Chondrosarcoma Cell Line CS8 - Grade 1
Chondrosarcoma
TAG LIB=UI-H-FHI
TAG_SEQ=AGAATCCGGC"

ORIGIN

Query Match 56.3%; Score 673.8; DB 5; Length 708;
Best Local Similarity 99.7%; Pred. No: 2.9e-175;
Matches 675; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 404 GGGTCAAAATAGAAATGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGA 463
DB 693 GGGTCAAAATAGAAATGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGA 634
QY 464 TTTCATTGACTTAGATCGTGTGGGCATCCACGGCTGGTCTTATGGAGGATACCTCTCCCT 523
DB 633 TTTCATTGACTTAGATCGTGTGGGCATCCACGGCTGGTCTTATGGAGGATACCTCTCCCT 574
QY 524 GATGGCATTAAATGCAGAGGTGATATCTTCAGGGTTCCTATTCGTGGGGCCCCAGTCA 583
DB 573 GATGGCATTAAATGCAGAGGTGATATCTTCAGGGTTCCTATTCGTGGGGCCCCAGTCA 514

QY 584 TCTGTGGATCTTCTATGATACAGGATACACGGAAGCTTATATGGGTACCTGACCCAGGAA 643
Db 513 TCTGTGGATCTTCTATGATACAGGATACACGGAAGCTTATATGGGTACCTGACCCAGGAA 454
QY 644 TGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTGAACC 703
Db 453 TGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTGAACC 394
QY 704 AATATCGTTTACTGCTTCTATGATACAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTGAACC 763
Db 393 AATATCGTTTACTGCTTCTATGATACAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTGAACC 334
QY 764 TATATTTACTAGTTTTTTAGTGGGCTGGAAGCCATATGATTTACAGATCTATCTCTCA 823
Db 333 TATATTTACTAGTTTTTTAGTGGGCTGGAAGCCATATGATTTACAGATCTATCTCTCA 274
QY 824 GGAGAGACAGCATTAAGAGTTTCTGAAATCGGGAGAACATTTATGAATGCTATCTTTTGA 883
Db 273 GGAGAGACAGCATTAAGAGTTTCTGAAATCGGGAGAACATTTATGAATGCTATCTTTTGA 214
QY 884 CTACCTTTCAAGAAACCTTTGGATCAGCTATTGCTGCTCTAAAGTGATATAATTTTGACC 943
Db 213 CTACCTTTCAAGAAACCTTTGGATCAGCTATTGCTGCTCTAAAGTGATATAATTTTGACC 154
QY 944 TGTGTAGAACTCTCTGATATACATCTGGCTATTAAACCAATGAGGAGGTTTAAATCAACAG 1003
Db 153 TGTGTAGAACTCTCTGATATACATCTGGCTATTAAACCAATGAGGAGGTTTAAATCAACAG 94
QY 1004 AAAACACAGAAATTTGATCATCATTTTGTATACCTGCCATGTAACATCTCTCTGAAAT 1063
Db 93 AAAACACAGAAATTTGATCATCATTTTGTATACCTGCCATGTAACATCTCTCTGAAAT 34
QY 1064 AATATGTTGGCCATGCA 1080
Db 33 AATATGTTGGCCATGNA 17

RESULT 15
AW303607/c
LOCUS
DEFINITION
xvzib2.x1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone
IMAGE:2813759 3' similar to TR:075273 075273 R26984_1 ;, mRNA
sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 673)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40UP from Gibco
High quality sequence stop: 465.
FEATURES
Location/Qualifiers
1..673
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="caxon:9606"
/clone IMAGE:2813759
/lab_hosts="DH10B"
/clone_libs="Soares_NFL_T_GBC_S1"
/note=Organ: pooled; Vector: pT73D-Pac (Pharmacia) with
a modified polylinker; Site 1: Not 1; Site 2: Eco RI;
Equal amounts of plasmid DNA from three normalized

libraries (fetal lung NbHL19W, testis NHT, and B-cell NCI-CGAP GCBI) were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as a tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of I.M.A.G.E. clones 297480-302087, 682632-687239, 726408-728711, and 729096-731399. Subtraction by Bento Soares and M. Fatima Bonaldo. "

ORIGIN

Query Match 55.9%; Score 668.8; DB 1; Length 673;
Best Local Similarity 99.6%; Pred. No. 7e-174;
Matches 670; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 510 GGATACCTCTCCCTGATGGCATTAAATGACAGAGGTGAGATATCTTCAGGGTTGCTATTGCT 569
Db 673 GGATACCTCTCCNTGATGGCATTAAATGACAGAGGTGAGATATCTTCAGGGTTGCTATTGCT 614
QY 570 GGGGCCCATGCTACTCTGTGGATCTTCTATGATACAGGATACACGGACGTTATATGGGT 629
Db 613 GGGGCCCATGCTACTCTGTGGATCTTCTATGATACAGGATACACGGACGTTATATGGGT 554
QY 630 CACCTTGACCAAGTGAACAGGGCTATTACTTAGGATCTGTGGCCTGACGAGCAAGAAAG 689
Db 553 CACCTTGACCAAGTGAACAGGGCTATTACTTAGGATCTGTGGCCTGACGAGCAAGAAAG 494
QY 690 TTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGTTTCTGGATGAGATGTCAT 749
Db 493 TTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGTTTCTGGATGAGATGTCAT 434
QY 750 TTTGCAATACACAGTATATTAATGAGTGTGAGGGCTGGAAGCCATATGATTTA 809
Db 433 TTTGCAATACACAGTATATTAATGAGTGTGAGGGCTGGAAGCCATATGATTTA 374
QY 810 CAGATCTATCTCAGAGAGACACAGCATAAGAGTTCTGAAATCGGAGAACATTATGAA 869
Db 373 CAGATCTATCTCAGAGAGACACAGCATAAGAGTTCTGAAATCGGAGAACATTATGAA 314
QY 870 CTGATCTTTTGCATCTACCTTCAAGAAACCTTTGGATCAGCTATTGCTGCTCTAAAAAGTG 929
Db 313 CTGATCTTTTGCATCTACCTTCAAGAAACCTTTGGATCAGCTATTGCTGCTCTAAAAAGTG 254
QY 930 ATATAATTTTGACCTGTGTAGAACTCTCTGGTATACACTGGCTATTTAAACCAATGAGGA 989
Db 253 ATATAATTTTGACCTGTGTAGAACTCTCTGGTATACACTGGCTATTTAAACCAATGAGGA 194
QY 990 GGTTTAATCAACAGAAACACAGAAATGATCATCACATTTTGTATACCTGCAATGTAACAT 1049
Db 193 GGTTTAATCAACAGAAACACAGAAATGATCATCACATTTTGTATACCTGCAATGTAACAT 134
QY 1050 CTACTCTGAAATAAATGTGTGCCATGAGGGGTCTACGGTTTGTGTAGTAATCTAA 1109
Db 133 CTACTCTGAAATAAATGTGTGCCATGAGGGGTCTACGGTTTGTGTAGTAATCTAA 74
QY 1110 TACCTTAACCCACATGCTCAAAATCAAAATGATATATTTCTGAGAGACCCAGCAATAC 1169
Db 73 TACCTTAACCCACATGCTCAAAATCAAAATGATATATTTCTGAGAGACCCAGCAATAC 14
QY 1170 CATAGAAATTAAT 1182
Db 13 CATAGAAATTAAT 1

Search completed: May 4, 2006, 07:56:18
Job time : 4493.98 secs

GenCore version 5.1.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 3, 2006, 01:01:26 ; Search time 378.893 Seconds
(without alignments)
5615.673 Million cell updates/sec

Title: US-10-825-632-4
Perfect score: 1197
Sequence: 1 attttgaaggcaccacaaagac.....ttactaaaaaaaaaaaaa 1197

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /cgn2_6/ptodata/1/ina/1 COMB.seq:*
- 2: /cgn2_6/ptodata/1/ina/5 COMB.seq:*
- 3: /cgn2_6/ptodata/1/ina/6A COMB.seq:*
- 4: /cgn2_6/ptodata/1/ina/6B COMB.seq:*
- 5: /cgn2_6/ptodata/1/ina/H COMB.seq:*
- 6: /cgn2_6/ptodata/1/ina/ECTUS COMB.seq:*
- 7: /cgn2_6/ptodata/1/ina/PP COMB.seq:*
- 8: /cgn2_6/ptodata/1/ina/RE COMB.seq:*
- 9: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1197	100.0	1197	3	US-10-070-464-4
2	1179	98.5	4676	3	US-09-976-674-20
3	1040	86.9	2797	3	US-09-976-594-1103
4	1036	86.5	3120	3	US-10-070-464-2
5	1033.2	86.3	4309	3	US-09-976-674-14
6	1033.2	86.3	4829	3	US-09-976-674-12
7	1013.6	84.7	4685	3	US-09-976-674-22
8	881	73.6	1669	3	US-10-070-464-6
9	873	72.9	4523	3	US-09-976-674-8
10	776.4	64.9	2671	3	US-09-976-674-2
11	633.4	52.9	823	3	US-09-280-116-171
12	459	38.3	1083	3	US-10-070-464-8
13	326.8	27.3	612	3	US-09-392-184-31
14	297	24.8	308	3	US-09-621-976-1787
15	212.8	17.8	2617	3	US-09-976-674-4
16	212.8	17.8	4180	3	US-09-976-674-36
17	212.8	17.8	4219	3	US-09-976-674-28
18	212.8	17.8	4263	3	US-09-976-674-34
19	212.8	17.8	4302	3	US-09-976-674-24
20	182.6	15.3	4076	3	US-09-976-674-32
21	182.6	15.3	4159	3	US-09-976-674-30
22	181.4	15.2	4037	3	US-09-976-674-40
23	181.4	15.2	4120	3	US-09-976-674-38
24	86	7.2	89	3	US-09-513-999C-16433

25	76.6	6.4	2946	3	US-09-902-540-9675	Sequence 9675, Ap
c	76.6	6.4	14555	3	US-09-902-540-1036	Sequence 1036, Ap
26	74.8	6.2	936	3	US-09-270-767-12940	Sequence 12940, A
27	73.4	6.1	1023	3	US-09-270-767-13726	Sequence 13726, A
28	70	5.8	3407	3	US-10-002-593-5	Sequence 5, Appli
29	70	5.8	3407	3	US-09-949-016-275	Sequence 275, App
30	70	5.8	3407	3	US-10-423-714-5	Sequence 5, Appli
31	70	5.8	3407	3	US-09-949-016-4579	Sequence 4579, Ap
32	68.4	5.7	2924	3	US-09-023-655-1026	Sequence 1026, Ap
33	68.4	5.7	2924	6	PCT-US93-07923-1	Sequence 1, Appli
34	65.4	5.5	2115	3	US-09-902-540-4518	Sequence 4518, Ap
35	65.4	5.5	20740	3	US-09-902-540-1223	Sequence 1223, Ap
36	64.4	5.4	2736	3	US-09-949-016-2459	Sequence 2459, Ap
37	64.4	5.4	2736	3	US-08-230-491A-1	Sequence 1, Appli
38	64.4	5.4	2815	2	US-08-619-280A-1	Sequence 1, Appli
39	64.4	5.4	2815	2	US-08-940-391-1	Sequence 1, Appli
40	64.4	5.4	2815	3	US-09-265-606-1	Sequence 1, Appli
41	61.8	5.2	535	3	US-09-389-681-428	Sequence 428, App
42	61.8	5.2	535	3	US-09-620-405B-428	Sequence 428, App
43	61.8	5.2	535	3	US-09-433-826B-428	Sequence 428, App
44	61.8	5.2	535	3	US-09-604-287A-428	Sequence 428, App
45	61.8	5.2	535	3		

ALIGNMENTS

RESULT 1

US-10-070-464-4
; Sequence 4, Application US/10070464
; Patent No. 6881564
; GENERAL INFORMATION:
; APPLICANT: ABBOTT, Catherine Anne
; APPLICANT: GORRELL, Mark Douglas
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES
; FILE REFERENCE: GH-007
; CURRENT APPLICATION NUMBER: US/10/070,464
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: PCT/AU00/01085
; PRIOR FILING DATE: 2000-09-11
; PRIOR APPLICATION NUMBER: AU PQ5709
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: AU PQ2762
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1197
; TYPE: DNA
; ORGANISM: Homo Sapiens
; US-10-070-464-4

Query Match	100.0%;	Score 1197;	DB 3;	Length 1197;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1197;	Conservative	0;	Mismatches	0;
Indels	0;	Gaps	0;	
QY	1	ATTTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTACGTAGTACGTAA	60	
Db	1	ATTTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTACGTAGTACGTAA	60	
QY	61	ATCCTTGAGAGGTGACAAAGGCTGACTGACCTGGCTACTCATCTTCTGTCATCAGTC	120	
Db	61	ATCCTTGAGAGGTGACAAAGGCTGACTGACCTGGCTACTCATCTTCTGTCATCAGTC	120	
QY	121	AGCACTGTGACTCTTTTATAGTAAGTATAGTAACAGAGAAATCCACACTGTGTGTC	180	
Db	121	AGCACTGTGACTCTTTTATAGTAAGTATAGTAACAGAGAAATCCACACTGTGTGTC	180	
QY	181	TTTCAAGCTATCAAGTCCTGAAGATGACCCAACTTGGCAAAACAAGAAATTTTGGCCA	240	
Db	181	TTTCAAGCTATCAAGTCCTGAAGATGACCCAACTTGGCAAAACAAGAAATTTTGGCCA	240	
QY	241	CCATTTTGGATTACAGAGGTCTCTTCTGCTACTATCTCTCCAGAAATTTTCTCTTTG	300	

Db 241 CCATTTTGGATTACAGAGTCTCTTCTCTGACTATACCTCCAGAAATTTTCTCTTTG 300
QY 301 AAAGTACTACTGGATTTACATTTGATGGATGCTCTACAGGCTCATGATCTACAGCTG 360
Db 301 AAAGTACTACTGGATTTACATTTGATGGATGCTCTACAGGCTCATGATCTACAGCTG 360
QY 361 GAAAGAAATATCTCTACTGCTCTTTCATATATGTTGGTCTCTCAGGCTCAATAGAAATG 420
Db 361 GAAAGAAATATCTCTACTGCTCTTTCATATATGTTGGTCTCTCAGGCTCAATAGAAATG 420
QY 421 ACGATCAGGTGGAGAGCTCCAAATATCTAGCTTCTCGATATGATTTCAATGACTTATG 480
Db 421 ACGATCAGGTGGAGAGCTCCAAATATCTAGCTTCTCGATATGATTTCAATGACTTATG 480
QY 481 GTGFGGCATCCACGGCTGGTCTATGAGGATACCTCTCCCTGATGCAATTAATGAGA 540
Db 481 GTGFGGCATCCACGGCTGGTCTATGAGGATACCTCTCCCTGATGCAATTAATGAGA 540
QY 541 GGTGAGATATCTTTGAGGTTGCTTATGCTGGGGCCCACTCACTCTGTGGATCTTCTATG 600
Db 541 GGTGAGATATCTTTGAGGTTGCTTATGCTGGGGCCCACTCACTCTGTGGATCTTCTATG 600
QY 601 ATACAGGATACAGGAACTTTATATGGTTCACCTGACGAGAAATGAACAGGGCTATTACT 660
Db 601 ATACAGGATACAGGAACTTTATATGGTTCACCTGACGAGAAATGAACAGGGCTATTACT 660
QY 661 TAGGATCTGTGGCATGCAAGAGAAAGTTCCTCTGAAACCAATCGTTTACTGCTCT 720
Db 661 TAGGATCTGTGGCATGCAAGAGAAAGTTCCTCTGAAACCAATCGTTTACTGCTCT 720
QY 721 TACATGTTTCTCTGGATGAGATGTCATTTTGCACATACCAATATATTAATGAGTTTTT 780
Db 721 TACATGTTTCTCTGGATGAGATGTCATTTTGCACATACCAATATATTAATGAGTTTTT 780
QY 781 TAGTGAGGCTGGAGAGCCATATGATTTACAGATCTATCTCCAGGAGACACAGCATAA 840
Db 781 TAGTGAGGCTGGAGAGCCATATGATTTACAGATCTATCTCCAGGAGACACAGCATAA 840
QY 841 GAGTTCCTGAATCGGAGACATATGATCTGATCTCTTTCGACTACCTTCAGAAACC 900
Db 841 GAGTTCCTGAATCGGAGACATATGATCTGATCTCTTTCGACTACCTTCAGAAACC 900
QY 901 TTGGATCAGTATTTGCTCTTAAAGATATATTTTGGACCTGTGTAGAACTCTCTGG 960
Db 901 TTGGATCAGTATTTGCTCTTAAAGATATATTTTGGACCTGTGTAGAACTCTCTGG 960
QY 961 TATACACTGGCTATTTAAACAAATGAGGAGGTTTAAATCAACAGAAACACAGAAATGATC 1020
Db 961 TATACACTGGCTATTTAAACAAATGAGGAGGTTTAAATCAACAGAAACACAGAAATGATC 1020
QY 1021 ATACATTTTGTATCTGCTGATGTAACATCTACTCTGAAATATAATGTGTGCTCATGCA 1080
Db 1021 ATACATTTTGTATCTGCTGATGTAACATCTACTCTGAAATATAATGTGTGCTCATGCA 1080
QY 1081 GGGGTCTACGGTTTGTGGTAGTAATCTAATACCTTTAACCCACATGCTCAAAATCAAAATG 1140
Db 1081 GGGGTCTACGGTTTGTGGTAGTAATCTAATACCTTTAACCCACATGCTCAAAATCAAAATG 1140
QY 1141 ATACATATTTCTGAGAGACCCAGCAATACCATAGAAATTAATAAATG 1197
Db 1141 ATACATATTTCTGAGAGACCCAGCAATACCATAGAAATTAATAAATG 1197

RESULT 2
US-09-976-674-20
; Sequence 20, Application US/09976674
; Patent No. 684180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV

; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; PILLING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PILLING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20
; LENGTH: 4676
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-976-674-20

Query Match 98.5%; Score 1179; DB 3; Length 4676;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1193; Conservative 0; Mismatches 0; Indels 4; Gaps 1;

QY 1 ATTTTGAAGGCACCAAGAGCTCCCTTTAGAGCATCACCTGTACGTAGTACGTAA 60
Db 1781 ATTTTGAAGGCACCAAGAGCTCCCTTTAGAGCATCACCTGTACGTAGTACGTAA 1840
QY 61 ATCTGTGAGAGGTGACAAAGGCTGACTGACCGTGGCTACTCACAATCTTGTGTCATCAGTC 120
Db 1841 ATCTGTGAGAGGTGACAAAGGCTGACTGACCGTGGCTACTCACAATCTTGTGTCATCAGTC 1900
QY 121 AGCACTGTGACTTCTTTTATAAGTAAGTATAGTAACAGAAAGATCCACACTGTGTGTC 180
Db 1901 AGCACTGTGACTTCTTTTATAAGTAAGTATAGTAACAGAAAGATCCACACTGTGTGTC 1960
QY 181 TTTTCAAGCTATCAAGTCTCTGAAGATGACCAACTTTGCAAAACAAAGGAATTTTGGGCA 240
Db 1961 TTTTCAAGCTATCAAGTCTCTGAAGATGACCAACTTTGCAAAACAAAGGAATTTTGGGCA 2020
QY 241 CCATTTTGAATTCAGCAGGCTCTCTTCTGACTATATCTCTCCAGAAATTTTCTTTG 300
Db 2021 CCATTTTGAATTCAGCAGGCTCTCTTCTGACTATATCTCTCCAGAAATTTTCTTTG 2080
QY 301 AAAGTACTACTGGATTTACATTTGATGGATGCTCTACAGGCTCATGATCTACAGCTG 360
Db 2081 AAAGTACTACTGGATTTACATTTGATGGATGCTCTACAGGCTCATGATCTACAGCTG 2140
QY 361 GAAAGAAATATCTCTACTGCTGCTTCAATATGTTGGTCTCTCAGGCTCAATAGAAATG 420
Db 2141 GAAAGAAATATCTCTACTGCTGCTTCAATATGTTGGTCTCTCAGGCTCAATAGAAATG 2196
QY 421 ACGATCAGGTGGAGAGCTCCAAATATCTAGCTTCTCGATATGATTTCAATGACTTAGATC 480
Db 2197 ACGATCAGGTGGAGAGCTCCAAATATCTAGCTTCTCGATATGATTTCAATGACTTAGATC 2256
QY 481 GTGFGGCATCCACGGCTGGTCTATGAGGATACCTCTCCCTGATGCAATTAATGAGCA 540
Db 2257 GTGFGGCATCCACGGCTGGTCTATGAGGATACCTCTCCCTGATGCAATTAATGAGCA 2316
QY 541 GGTGAGATATCTTTGAGGTTGCTTATGCTGGGGCCCACTCACTCTGTGGATCTTCTATG 600
Db 2317 GGTGAGATATCTTTGAGGTTGCTTATGCTGGGGCCCACTCACTCTGTGGATCTTCTATG 2376
QY 601 ATACAGGATACAGGAACTTTATATGGTTCACCTGACGAGAAATGAACAGGGCTATTACT 660
Db 2377 ATACAGGATACAGGAACTTTATATGGTTCACCTGACGAGAAATGAACAGGGCTATTACT 2436
QY 661 TAGGATCTGTGGCATGCAAGAGAAAGTTCCTCTGAAACCAATCGTTTACTGCTCT 720
Db 2437 TAGGATCTGTGGCATGCAAGAGAAAGTTCCTCTGAAACCAATCGTTTACTGCTCT 2496
QY 721 TACATGTTTCTCTGGATGAGAAATGTCATTTTGCACATACCAATATATTAATGAGTTTTT 780
Db 2497 TACATGTTTCTCTGGATGAGAAATGTCATTTTGCACATACCAATATATTAATGAGTTTTT 2556
QY 781 TAGTGAGGCTGGAGAGCCATATGATTTTACAGATCTATCTCCAGGAGACACAGCATAA 840
Db 2557 TAGTGAGGCTGGAGAGCCATATGATTTTACAGATCTATCTCCAGGAGACACAGCATAA 2616

QY	514	ACCTCTCCCTGATGGCAATTAATGCAGAGTCAGATATCTTTCAGGGTGCTATTCTCGTGGG	573
DB	2441	ACCTCTCCCTGATGGCAATTAATGCAGAGTCAGATATCTTTCAGGGTGCTATTCTCGTGGG	2500
QY	574	CCCAGTCACCTCTGTGGATCTTATGATACAGGATACACGGAACTTTATATGGSTCACC	633
DB	2501	CCCAGTCACCTCTGTGGATCTTATGATACAGGATACACGGAACTTTATATGGSTCACC	2560
QY	634	CTGACCAGATGAACAGGGCTATTACTTAGGATCTGTGGCATGCAAGCAGAAAAGTTCC	693
DB	2561	CTGACCAGATGAACAGGGCTATTACTTAGGATCTGTGGCATGCAAGCAGAAAAGTTCC	2620
QY	694	CCTCTGAAACCAAATCGTTTTACTGCTCTTACATGGTTTTCTTGGATGAGAATGCCATTTTG	753
DB	2621	CCTCTGAAACCAAATCGTTTTACTGCTCTTACATGGTTTTCTTGGATGAGAATGCCATTTTG	2680
QY	754	CACATPACCAGPATATACTAGATTTTTTATGTAGGGCTGGAAAGCCATATGATTTACGA	813
DB	2681	CACATPACCAGPATATACTAGATTTTTTATGTAGGGCTGGAAAGCCATATGATTTACGA	2740
QY	814	TCTATCCTCAGGAGAGACACAGCATAAGAGTTCTCTGAAATCGGGAGAACATTATGAATGTC	873
DB	2741	TCTATCCTCAGGAGAGACACAGCATAAGAGTTCTCTGAAATCGGGAGAACATTATGAATGTC	2800
QY	874	ATCTTTTTGCACCTACCTTTCAAGAAAAACCTTTGGATCACGTATTGCTCTCTAAAGTGATAT	933
DB	2801	ATCTTTTTGCACCTACCTTTCAAGAAAAACCTTTGGATCACGTATTGCTCTCTAAAGTGATAT	2860
QY	934	AATTTTGGACCTGTGTAGAACTCTCTGGTATACACTGGGCTATTATAACCAAAATGAGGAGTT	993
DB	2861	AATTTTGGACCTGTGTAGAACTCTCTGGTATACACTGGGCTATTATAACCAAAATGAGGAGTT	2920
QY	994	TAATCAACAGAAAAACACAGAAATTTGATCATCACATTTTGATACCTGCCATGTAACATCTAC	1053
DB	2921	TAATCAACAGAAAAACACAGAAATTTGATCATCACATTTTGATACCTGCCATGTAACATCTAC	2980
QY	1054	TCCTGAAAATAAAATGTGGTGCCATGCAGGGGCTACGGTTTCTGGTAGTAATCTAAATACC	1113
DB	2981	TCCTGAAAATAAAATGTGGTGCCATGCAGGGGCTACGGTTTCTGGTAGTAATCTAAATACC	3040
QY	1114	TTAAACCCACATGCTCAAAATCAAATGATACATATTTCTTGAGAGAGCCAGCAATACCATA	1173
DB	3041	TTAAACCCACATGCTCAAAATCAAATGATACATATTTCTTGAGAGAGCCAGCAATACCATA	3100
QY	1174	AGAAATTACTTAAAAAANA 1193	
DB	3101	AGAAATTACTTAAAAAANA 3120	
 RESULT 5 US-09-976-674-14 ; Sequence 14, Application US/09976674 ; Patent No. 6844180 ; GENERAL INFORMATION: ; APPLICANT: Qi, Steve ; APPLICANT: Akinsanya, Karen ; APPLICANT: Riviere, Pierre ; APPLICANT: Junien, Jean-Louis ; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV ; FILE REFERENCE: 70669 ; CURRENT APPLICATION NUMBER: US/09/976, 674 ; CURRENT FILING DATE: 2001-10-12 ; PRIOR APPLICATION NUMBER: US 60/240,117 ; PRIOR FILING DATE: 2000-10-12 ; NUMBER OF SEQ ID NOS: 61 ; SOFTWARE: PatentIn version 3.1 ; SEQ ID NO 14 ; LENGTH: 4309 ; TYPE: DNA ; ORGANISM: Homo sapiens US-09-976-674-14			

Db	2281	GCATCTTTTGGCACTACCTTTCAAGAAACCTTGGATCAGGTATTGCTGCTCTTAAAGATGAT	2344
Qy	932	ATAATTTTGAAGCTGTGTAGAACTCTCTGCTATATACACTGGGTATTTAAACCAATGAGGAGG	991
Db	2341	ATAATTTTGAAGCTGTGTAGAACTCTCTGCTATATACACTGGGTATTTAAACCAATGAGGAGG	2400
Qy	992	TTTAAATCAACAGAAACACAGAAATGATCATCATTTTGTATACCTGCCATGTAACATCT	1051
Db	2401	TTTAAATCAACAGAAACACAGAAATGATCATCATTTTGTATACCTGCCATGTAACATCT	2460
Qy	1052	ACTCTCTGAAATTAATGTGGTCCCATGCAGGGGTCTACGGTTTCTGTGTAGTAACTTAATA	1111
Db	2461	ACTCTCTGAAATTAATGTGGTCCCATGCAGGGGTCTACGGTTTCTGTGTAGTAACTTAATA	2520
Qy	1112	CCTTAACCCCATCATCTCTCAAAATCAAAATGATACATATTTCTTGAGAGACCCAGCAATACCA	1171
Db	2521	CCTTAACCCCATCATCTCTCAAAATCAAAATGATACATATTTCTTGAGAGACCCAGCAATACCA	2580
Qy	1172	TAAGAATTACTAAAAAATAAAAAAAAAAAAAA	1197
Db	2581	TAAGAATTACTAAAAAATAAAAAAAAAAAAAA	2606
RESULT 6			
US-09-976-674-12			
; Sequence 12, Application US/09976674			
; Patent No. 6844180			
; GENERAL INFORMATION:			
; APPLICANT: Akinsanya, Karen			
; APPLICANT: Riviere, Pierre			
; APPLICANT: Junien, Jean-Louis			
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV			
; FILE REFERENCE: 70669			
; CURRENT APPLICATION NUMBER: US/09/976,674			
; CURRENT FILING DATE: 2001-10-12			
; PRIOR FILING DATE: 2001-10-12			
; NUMBER OF SEQ ID NOS: 61			
; SOFTWARE: PatentIn version 3.1			
; SEQ ID NO 12			
; LENGTH: 4829			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
US-09-976-674-12			
Query Match 86.3%; Score 1033.2; DB 3; Length 4829;			
Best Local Similarity 88.7%; Pred. No: 0;			
Matches 1194; Conservative 0; Mismatches 3; Indels 149; Gaps 1;			
Qy	1	ATTTTGAAGGCACCAAGACTCCCTTTAGAGCATCCTCTAGAGTACAGTACGTAA	60
Db	1781	ATTTTGAAGGCACCAAGACTCCCTTTAGAGCATCCTCTAGAGTACAGTACGTAA	1840
Qy	61	ATCCTGGAGGTGACAGGGCTGACTGACCGTGGCTACTCACATTTCTGCTGCATCAGTC	120
Db	1841	ATCCTGGAGGTGACAGGGCTGACTGACCGTGGCTACTCACATTTCTGCTGCATCAGTC	1900
Qy	121	AGCACTGTGACTTCTTTATAGTAAGTATAGTAACCAAGAAATCCACACTGTGTGCC	180
Db	1901	AGCACTGTGACTTCTTTATAGTAAGTATAGTAACCAAGAAATCCACACTGTGTGCC	1960
Qy	181	TTTACAAGCTATCAAGTCTCTGAAGATGACCAACTTGCACAAAGAAATTTTGGGCCA	240
Db	1961	TTTACAAGCTATCAAGTCTCTGAAGATGACCAACTTGCACAAAGAAATTTTGGGCCA	2020
Qy	241	GCATTTTGAATTCAGCAGGTCTCTCTCTGACTATATCTCTCCAGAAATTTTCTTTTG	300
Db	2021	GCATTTTGAATTCAGCAGGTCTCTCTCTGACTATATCTCTCCAGAAATTTTCTTTTG	2080
Qy	301	AAAGTACTACTGGATTTTACATTTGATGGATGCTCTACAGCCTCATGATCTACAGCCTG	360
Db	2081	AAAGTACTACTGGATTTTACATTTGATGGATGCTCTACAGCCTCATGATCTACAGCCTG	2140

361 GAAAGAAATATCTACTGTGCTGTTTCATATATGTTGGTCTC----- 402
Db |||||
2141 GAAAGAAATATCTACTGTGCTGTTTCATATATGTTGGTCTCCTCAGGTGCAAGTTGGTGAA 2200
Qy |||||
403 ----- 402
Db
2201 TAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTAGCCTCTCTAGGTTATGT 2260
Qy ----- 402
Db
2261 GGTGTAGTATAGACACAGAGGGATCCGTCCAGAGGGCTTAAATTTGAAGCGCGCTT 2320
Qy ----- 403
Db
403 -----AGGTTCAATAGAAATTTGACGATCAGGTGGAAGACTCCAAATATCTPAGC 451
Db
2321 TAAATATAAATGGGTCAAAATGAAATTTGACGATCAGGTGGAAGACTCCAAATATCTPAGC 2380
Qy
452 TTCTCGATATGATTTTCAATTTAGACTTATGATCGTGTGGGCATCCACGGCTGGTCCCTATGGAGG 511
Db
2381 TTCTCGATATGATTTTCAATTTAGACTTATGATCGTGTGGGCATCCACGGCTGGTCCCTATGGAGG 2440
Qy
512 ATACCTCTCCCTGATGGCATTAAATGCAGAGGTTCAGATATCTTTCAGGGTTGCTATTGCTGG 571
Db
2441 ATACCTCTCCCTGATGGCATTAAATGCAGAGGTTCAGATATCTTTCAGGGTTGCTATTGCTGG 2500
Qy
572 GGCCCCAGTCACTCTGTGATCTTCTATGATACAGGATACACGGAACTTTATATGGGTCA 631
Db
2501 GGCCCCAGTCACTCTGTGATCTTCTATGATACAGGATACACGGAACTTTATATGGGTCA 2560
Qy
632 CCTGTACAGAAATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAAGTT 691
Db
2561 CCTGTACAGAAATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAAGTT 2620
Qy
692 CCCTCTGAACCAATCGTTTACTGCTCTTACATGTTTCTCGATGAGATGCTCCATTT 751
Db
2621 CCCTCTGAACCAATCGTTTACTGCTCTTACATGTTTCTCGATGAGATGCTCCATTT 2680
Qy
752 TGCAATACAGTATATTTACTGCTGTTTCTAGTGGCTGGAAGCCATATGATTTACA 811
Db
2681 TGCAATACAGTATATTTACTGCTGTTTCTAGTGGCTGGAAGCCATATGATTTACA 2740
Qy
812 GATCTATCTCAGGAGACACAGATTAAGGTTTCTGATCGGAGAACATPATGAAT 871
Db
2741 GATCTATCTCAGGAGACACAGATTAAGGTTTCTGATCGGAGAACATPATGAAT 2800
Qy
872 GATCTTTTGCATACCTTCAAGAAACCTTTGGATCAGTATTGCTGCTCTAAAGTGAT 931
Db
2801 GATCTTTTGCATACCTTCAAGAAACCTTTGGATCAGTATTGCTGCTCTAAAGTGAT 2860
Qy
932 ATAAATTTGACCTGTGTAGAACTCTCTGGTATACACTGGCTATTAAACCAAAATGAGGAGG 991
Db
2861 ATAAATTTGACCTGTGTAGAACTCTCTGGTATACACTGGCTATTAAACCAAAATGAGGAGG 2920
Qy
992 TTTAATCAACAGAAACACAGAAATTTGATCATCATTTTGGATACCTGCTGCTTAAATCT 1051
Db
2921 TTTAATCAACAGAAACACAGAAATTTGATCATCATTTTGGATACCTGCTGCTTAAATCT 2980
Qy
1052 ACTCTGAAATTAATTTGGTGCCATGAGGGGCTACGGTTTGGTGTAGTAAATCTAATA 1111
Db
2981 ACTCTGAAATTAATTTGGTGCCATGAGGGGCTACGGTTTGGTGTAGTAAATCTAATA 3040
Qy
1112 CTTTAAACCCCATCTCTCAAAATCAAAATGATACATATCTCTGAGAGACCCAGCAATACCA 1171
Db
3041 CTTTAAACCCCATCTCTCAAAATCAAAATGATACATATCTCTGAGAGACCCAGCAATACCA 3100
Qy
1172 TAAGATTAATAAAAAAAAAAAAAA 1197
Db
3101 TAAGATTAATAAAAAAAAAAAAAA 3126

RESULT 7
US-09-976-674-22
; Sequence 22, Application US/09976674

; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 22
; LENGTH: 4685
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-976-674-22

Query Match 84.7%; Score 1013.6; DB 3; Length 4685;
Best Local Similarity 91.3%; Pred. No. 0;
Matches 1098; Conservative 0; Mismatches 99; Indels 5; Gaps 2;

Qy 1 ATTTGAAGGCACCAAGAGACTCCCTTTAGAGCATCCTGTAGTACTCAGTTACGTAA 60
Db 1781 ATTTGAAGGCACCAAGAGACTCCCTTTAGAGCATCCTGTAGTACTCAGTTACGTAA 1840
Qy 61 ATCTGGAGGTGACCAAGGCTGACTGACCGTGGCTACTACATCTTGTGATCAGTC 120
Db 1841 ATCTGGAGGTGACCAAGGCTGACTGACCGTGGCTACTACATCTTGTGATCAGTC 1900
Qy 121 AGCACTGTACATCTTTTAAAGTATAGTAACTGACCAAGAAATCCACATCTGTGTCCC 180
Db 1901 AGCACTGTACATCTTTTAAAGTATAGTAACTGACCAAGAAATCCACATCTGTGTCCC 1960
Qy 181 TTTTCAAGCTATCAAGTCTCTGAAGATGACCAACTTTGCAAAACAAAGAAATTTTGGGCCA 240
Db 1961 TTTTCAAGCTATCAAGTCTCTGAAGATGACCAACTTTGCAAAACAAAGAAATTTTGGGCCA 2020
Qy 241 CCATTTTGGATTGAG --- CAGGTCTCTTCTCTGACT -ATACTCTCCAGAAATTTTCTC 295
Db 2021 CCATTTTGGATTGAGTTTCTCTCAGGTGCTGTTGAATAATCGGTTTAAAGAGTCAAGTA 2080
Qy 296 TTTTGAAGTACTACTGGATTTTACATTTGTATGGGATGCTCTCAAGGCTCATGATCTACA 355
Db 2081 TTTCCGCTTGAATACCTAGCCTCTCTAGTTTATGTTGTTAGTGATAGACACAGGGG 2140
Qy 356 GCCTGAAAGAAATATCTCTACTGTCTGTTTATATATGTTGGTCTCCTCAGGGTCAAAATAGA 415
Db 2141 ATCTCTCACCGAGGCTTAAATTTGAAGGCGCCTTTAAATATAAAATGGGTCAAAATAGA 2200
Qy 416 AATTGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTTCAATTTGACTT 475
Db 2201 AATTGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTTCAATTTGACTT 2260
Qy 476 AGATCGTGTGGCATCCACGGCTGCTTATGAGGATACCTCTCCCTGATGGCATTAAT 535
Db 2261 AGATCGTGTGGCATCCACGGCTGCTTATGAGGATACCTCTCCCTGATGGCATTAAT 2320
Qy 536 GCAGAGGTGATATCTTTCAGGGTGTCTTTCAGTGGGCCCCAGTCTCTGTGGATCTT 595
Db 2321 GCAGAGGTGATATCTTTCAGGGTGTCTTTCAGTGGGCCCCAGTCTCTGTGGATCTT 2380
Qy 596 CTATGATACAGGATACAGGACGTTATATGGGTCACTGACCAAGCAATGAACAGGGCTA 655
Db 2381 CTATGATACAGGATACAGGACGTTATATGGGTCACTGACCAAGCAATGAACAGGGCTA 2440
Qy 656 TTACTTAGGATCTGTGGCCATGCAAGCAAGAAAGTTTCCCTCTGAAACAAATCGTTTACT 715
Db 2441 TTACTTAGGATCTGTGGCCATGCAAGCAAGAAAGTTTCCCTCTGAAACAAATCGTTTACT 2500
Qy 716 GCTCTTACATGGTTTCTCTGGATGAGAAATGTCCTCAATTTTGCATACCAAGTATATTTACTGAG 775


```
QY 181 TTACAACTATCAAGTCTGAGATGACCAACTTGGCAAAACAAAGGAATTTGGGCA 240
Db 1755 TTACAACTATCAAGTCTGAGATGACCAACTTGGCAAAACAAAGGAATTTGGGCA 1814
QY 241 CCAATTTGGATTCAGCAGGCTCTCTCTGACTATATCTCTCCAGAAAATTTCTCTTTG 300
Db 1815 CCAATTTGGATTCAGCAGGCTCTCTCTGACTATATCTCTCCAGAAAATTTCTCTTTG 1874
QY 301 AAAGTACTACTGGATTTACATTTGATGGGATGCTCTACAGGCTCATGATCTACAGCTG 360
Db 1875 AAAGTACTACTGGATTTACATTTGATGGGATGCTCTACAGGCTCATGATCTACAGCTG 1934
QY 361 GAAAGAAATATCTACTGCTGCTTTCATATATGCTGCTCA----- 403
Db 1935 GAAAGAAATATCTACTGCTGCTTTCATATATGCTGCTCAAGGTGGTGAATA 1994
QY 404 ----- 403
Db 1995 ATCGATTTAAGGAGTCAAGTATTTCCGCTTGAATACCTAGCTCTCTAGGTATGTGG 2054
QY 404 ----- 403
Db 2055 TTGTAGTAGATAGACACAGGGATCTGTCCAGCGGGCTTAAATTTGAAGCGCTTTA 2114
QY 404 -----GGGTCAAATAGAAATTAAGATGAGATGAGGATCTCAATATCTAGCTT 453
Db 2115 AATATAAAATGGGTCAAATAGAAATTAAGATGAGATGAGGATCTCAATATCTAGCTT 2174
QY 454 CTCGATGATTTCAATTCATGACTTAGATGCTGTGGGATCCAGGCTGCTCTATGAGGAT 513
Db 2175 CTCGATGATTTCAATTCATGACTTAGATGCTGTGGGATCCAGGCTGCTCTATGAGGAT 2234
QY 514 ACCTCTCCCTGATGGCTTAATGCAGAGGTCAGATATCTTCAGGCTGCTATTCCTGGG 573
Db 2235 ACCTCTCCCTGATGGCTTAATGCAGAGGTCAGATATCTTCAGGCTGCTATTCCTGGG 2294
QY 574 CCCAGTCACCTCTGTGGATCTTCTATGATACAGGATACACGGAAGTTATATGGTCAAC 633
Db 2295 CCCAGTCACCTCTGTGGATCTTCTATGATACAGGATACACGGAAGTTATATGGTCAAC 2354
QY 634 CTGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCTCAAGCAGAAAAGTTCC 693
Db 2355 CTGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCTCAAGCAGAAAAGTTCC 2414
QY 694 CCTCTGAACCAAAATCGTTTACTGCTCTTACATGTTTCTGATGAGAAATGTCATTTG 753
Db 2415 CCTCTGAACCAAAATCGTTTACTGCTCTTACATGTTTCTGATGAGAAATGTCATTTG 2474
QY 754 CACATACAGTATATTAATGAGTTTTTATGAGGCTGGAAGCCATATGATTTACAGA 813
Db 2475 CACATACAGTATATTAATGAGTTTTTATGAGGCTGGAAGCCATATGATTTACAGA 2534
QY 814 TCTATCTCAGGAGACACAGGATAGAGTTCTGAAATCGGGGAACATTTAGACTGC 873
Db 2535 TCTATCTCAGGAGACACAGGATAGAGTTCTGAAATCGGGGAACATTTAGACTGC 2594
QY 874 ATCTTTTGCACTACTCTTCAAGAAAACCTTGGATCAGTATTCCTGCTCTAAAAGTGATAT 933
Db 2595 ATCTTTTGCACTACTCTTCAAGAAAACCTTGGATCAGTATTCCTGCTCTAAAAGTGATAT 2654
QY 934 AA 935
Db 2655 GA 2656
```

RESULT 11

US-09-280-116-171
; Sequence 171, Application US/09280116A
; Patent No. 6331427
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: Nucleic Acid Molecules Encoding Human Protease Homologs

FILE REFERENCE: 5800-24, 035800/176965
; CURRENT APPLICATION NUMBER: US/09/280,116A
; CURRENT FILING DATE: 1999-03-26
; NUMBER OF SEQ ID NOS: 268
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 171
; LENGTH: 823
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: prolyl oligopeptidases
US-09-280-116-171

Query Match 52.98; Score 633.4; DB 3; Length 823;
Best Local Similarity 94.3%; Pred. No. 1.1e-196;
Matches 776; Conservative 0; Mismatches 11; Indels 36; Gaps 10;
QY 406 GTCAAAATAGAATTTGACGATCAGGTGGAAGGAC-TCCAATATCTAGCTTCTCGATATGAT 464
Db 1 GTCAAAATAGAATTTGACGATCAGGTGGAAGGACATCCAATATCTAGCTTCTCGATATGAT 60
QY 465 TTCATTGACTTAGATTCGTGTGGCATCCAGGCTGCTCTATGAGGATACCTCTCCCTG 524
Db 61 TTCATTGACTTAGATTCGTGTGGCATCCAGGCTGCTCTATGAGGATACCTCTCCCTG 120
QY 525 ATGGCATTAATGACAGAGTTCAGATATCTTCAGGTTGCTTATGCTGGGGCCCCAGTCACT 584
Db 121 ATGGCATTAATGACAGAGTTCAGATATCTTCAGGTTGCTTATGCTGGGGCCCCAGTCACT 180
QY 585 CTGTGGATCTTCTATGATACAGGATACACGGAACGTTTATATGGTCAACCTGACCAAGAT 644
Db 181 CTGTGGATCTTCTATGATACAGGATACACGGAACGTTTATATGGTCAACCTGACCAAGAT 240
QY 645 GAAACAGGGCTATTACTTAGGATCTGTGGCATCAGCAGAGAAAAGTTCCCTCTGAACA 704
Db 241 GAAACAGGGCTATTACTTAGGATCTGTGGCATCAGCAGAGAAAAGTTCCCTCTGAACA 300
QY 705 AATCGTTTACTGCTCTTACATGTTTCTCGATGAGAAATGTCATTTGTCATACCACT 764
Db 301 AATCGTTTACTGCTCTTACATGTTTCTCGATGAGAAATGTCATTTGTCATACCACT 360
QY 765 ATATTACTAGTCTTTTGTAGGCTGGAAGGCCATATGATTTAC----- 810
Db 361 ATATTACTAGTCTTTTGTAGGCTGGAAGGCCATATGATTTTACAGTATCTTTATTTT 420
QY 811 -----AGATCTATCTCTCA-GGAGAGACA-CAGCATAGAGTTTCCG--AATCG 854
Db 421 TTGTTTGTGTAAGATCTATCTCTCAGGAGAGACACCGGCATAAGAGTTCCCTGAATCG 480
QY 855 GGAGAACATTTATGAACTGCACTTTT-GACTACCTTCAAG-AAACCTTGGATCA-CGT 911
Db 481 GGAGAACATTTATGAACTGCACTTTTGGCACTACCTTCAAGAAAAACCTTGGATCACCTT 540
QY 912 ATTCGCTCT-CTAAAAGTGATATAATTTGACCTGTGTAGAACTCTCT-GGTATACACTG 969
Db 541 ATTCGCTCTCTAAAAGTGATATAATTTGACCTGTGTAGAACTCTCTGGGGTATACACTG 600
QY 970 GCTATTTAAACCAATAGGAGGTTTAAATGAACAGAAAAACACAGAAATGATCATCATTT 1029
Db 601 GCTATTTAAACCAATAGGAGGTTTAAATGAACAGAAAAACACAGAAATGATCATCATTT 660
QY 1030 TGATACCTGTCATGTAACATCTACTCTCTGAAAAATAAATGTGTGGCTGCGAGGGGTCTAC 1089
Db 661 TGATACCTGTCATGTAACATCTACTCTCTGAAAAATAAATGTGTGGCTGCGAGGGGTCTAC 720
QY 1090 GGTGTGTGTAGTAATCTAAATACCTTAAGCCCAATGCTCAAAATCAAAATGATACATATT 1149
Db 721 GGTGTGTGTAGTAATCTAAATACCTTAAGCCCAATGCTCAAAATCAAAATGATACATATT 780
QY 1150 CTTGAGAGACCCAGCAATACCATAGAAATTTACTTAAAAAATAA 1192
Db 781 CTTGAGAGACCCAGCAATACCATAGAAATTTACTTAAAAAATAA 823

```
RESULT 12
US-10-070-464-8
; Sequence 8, Application US/10070464
; Patent No. 6881564
; GENERAL INFORMATION:
; APPLICANT: ABBOTT, Catherine Anne
; APPLICANT: GORRELL, Mark Douglas
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES
; FILE REFERENCE: GH-007
; CURRENT APPLICATION NUMBER: US/10/070,464
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: PCT/AU00/01085
; PRIOR FILING DATE: 2000-09-11
; PRIOR APPLICATION NUMBER: AU PQ5709
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: AU PQ2762
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 1083
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-10-070-464-8

Query Match      38.3%; Score 459; DB 3; Length 1083;
Best Local Similarity 80.7%; Pred. No. 1.8e-139;
Matches 616; Conservative 0; Mismatches 0; Indels 147; Gaps 1;

QY 1 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACCTGTACGTAGTCAGTTACGTAA 60
DB 321 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACCTGTACGTAGTCAGTTACGTAA 380

QY 61 ATCTGTGAGAGGTGACAAAGCTGACTGACCGTGGCTACTCACTTCTTGTGCAATCAGTC 120
DB 381 ATCTGTGAGAGGTGACAAAGCTGACTGACCGTGGCTACTCACTTCTTGTGCAATCAGTC 440

QY 121 AGCACTGTGACTTCTTTATAAGTAAGTATAGTAAGTAAACAGAAAGTCCCACTGTGTGCC 180
DB 441 AGCACTGTGACTTCTTTATAAGTAAGTATAGTAAGTAAACAGAAAGTCCCACTGTGTGCC 500

QY 181 TTTTACAAGCTATCAAGTCTCTGAAGATGACCCAACTTGGCAAAACAAAGAAATTTTGGGCCA 240
DB 501 TTTTACAAGCTATCAAGTCTCTGAAGATGACCCAACTTGGCAAAACAAAGAAATTTTGGGCCA 560

QY 241 CCATTTTGGATTACAGAGGTCTCTTCTGCTACTATCTCTCCAGAAATTTCTCTTTTG 300
DB 561 CCATTTTGGATTACAGAGGTCTCTTCTGCTACTATCTCTCCAGAAATTTCTCTTTTG 620

QY 301 AAAGTACTACTGGATTACATTTGTATGGATGCTCTACAAGCTCATGATCAGAGCTG 360
DB 621 AAAGTACTACTGGATTACATTTGTATGGATGCTCTACAAGCTCATGATCAGAGCTG 680

QY 361 GAAAGAAATATCCTACTGTGCTGTTTCATATATGTTGGTCTCTCA----- 403
DB 681 GAAAGAAATATCCTACTGTGCTGTTTCATATATGTTGGTCTCTCA----- 740

QY 404 ----- 403
DB 741 ATCGGTTTAAAGGAGTCAAGATATTTCCGCTTGAATACCTAGACCTCTCTAGGTTATGTGG 800

QY 404 ----- 403
DB 801 TTGTAGTGATAGACACAGGGATCCTGTCAACCGAGGCTTAAATTTTGAAGGCGCCCTTTA 860

QY 404 -----GGGTCAAATAGAAATTTGACGATCAGGTGGAAGGACTCCAAATATCTAGCTTT 453
DB 861 AATATAAAATGGGTCAAATAGAAATTTGACGATCAGGTGGAAGGACTCCAAATATCTAGCTTT 920

QY 454 CTCGATATGATTTCAITGATTAGATCGTGTGGGATCCAGGGTGGTCTCTATGGAGGAT 513
DB 921 CTCGATATGATTTCAITGATTAGATCGTGTGGGATCCAGGGTGGTCTCTATGGAGGAT 980
```

```
QY 514 ACCTCTCCCTGATGGCAATTAATGCAGAGGTACAGATATCTTCAGGGTGTCTATTCTCGGG 573
DB 981 ACCTCTCCCTGATGGCAATTAATGCAGAGGTACAGATATCTTCAGGGTGTCTATTCTCGGG 1040

QY 574 CCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGA 616
DB 1041 CCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGA 1083

RESULT 13
US-09-392-184-31/c
; Sequence 31, Application US/09392184
; Patent No. 6395889
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN
; FILE REFERENCE: PROTEASE HOMOLOGS
; FILE REFERENCE: 5800-55
; CURRENT APPLICATION NUMBER: US/09/392,184
; CURRENT FILING DATE: 1999-09-09
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 31
; LENGTH: 612
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(612)
; OTHER INFORMATION: prolyloligo (prolyl oligopeptidase)
; NAME/KEY: misc feature
; LOCATION: (1)..(612)
; OTHER INFORMATION: n = A,T,C or G
US-09-392-184-31

Query Match      27.3%; Score 326.8; DB 3; Length 612;
Best Local Similarity 79.2%; Pred. No. 2.6e-96;
Matches 397; Conservative 0; Mismatches 101; Indels 3; Gaps 3;

QY 413 AGAAATTTGACGATCAGGTGGAAGGACT-CCAAATATCTAGCTTCTCGATATGATTTCAATTG 471
DB 520 AGAATTTGCGATCAGGTGGNGGACTCCCAATATCTAACTTCTCGAAATGATTTCAATTG 461

QY 472 ACTTAGATCTGTGGGCTCCACGGCTGCTCTATGGAGGATACCTCTCCCTGATGGCAT 531
DB 460 ACNTAAGATCTGTGGCAATCCNCGGCTGCTCTATGGAGGATACCTCTCCCTGATGGCAN 401

QY 532 TAATGACAGAGGTGAGATATCTTCAGGGTTGCTATTGCTGGGGCCCCAGTCACTCTGTGGA 591
DB 400 TAATTCAGAGGTGAGATATCTTCAGGGTT-CTATTTCTGGGGCCCCAGTCACTCTGTGGA 342

QY 592 TCTTCTATGATACAGGATACAGGACGTTATATGGTCACTTGGTCACTTNNCCAGNATGGACAGG 651
DB 341 TCTTCTATGAAACAGGAAACNCGGACGTTANATGGTCACTTNNCCAGNATGGACAGG 282

QY 652 GCTATTTACTTAGATCTGTGGCATCGCAGGAAAGTTCCTCTGGAACCAATCGTT 711
DB 281 GCTATTTACTTAGATCTGTGGCATTCANAGGAAAGTTCCTCTGNNCCCAATNNGT 222

QY 712 TACTGTCTTACATGTTTCTCGGATGAGAAATGTCATTTTGCACATACCATATATAC 771
DB 221 TNCCTCTCTANATGTTTCTCGGATTAGAATTTTCANATTTTCANACCAAGNANATAC 162

QY 772 TGAGTTTTTTAGTGAGGGCTGGAAGCCATATGATTTTACATCTATCTCTCAGAGAGAC 831
DB 161 TNAGTTTTTTAGTGAGGGCTGGAAGCCAAATGANTTNCAGAACCACTCNGAGAGAGN 102

QY 832 ACAGCAATAGAGTTCTGAAATCGGAGAGAAATATATG-AACTGCAATCTTTTGCATACCTT 890
DB 101 CCAACANNAGAGTCTGAAATCGGAGGAGCAATATNTGACCNNTNACCTTTTNNACCCCN 42

QY 891 CAAGAAAAACCTTGGATCAGCT 911
```

Db 41 CAGAAACCTTGGATCACNT 21
||||| ||||||| |||||||

RESULT 14

US-09-621-976-1787
; Sequence 1787, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET 054PE2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 1787
; LENGTH: 308
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 17..178
; NAME/KEY: sig_peptide
; LOCATION: 17..133
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 4.59999990463257
; OTHER INFORMATION: seq AHTSILLFLVRA/GK
US-09-621-976-1787

Query Match 24.8%; Score 297; DB 3; Length 308;

Best Local Similarity 99.7%; Pred. No. 9.9e-87; Mismatches 0; Indels 1; Gaps 1;

Matches 308; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 659 CTTAGGATCTGTGGCCATGCAAGCAGAAAAAGTTCCCTCTGAACCAAAATCGTTTACTGCT 718
Db 1 CTTAGGATCTGTGGCCATGCAAGCAGAAAAAGTTCCCTCTGAACCAAAATCGTTTACTGCT 60

QY 719 CTTACATGGTTTCTCGATGAGAAATGTCATTTGGACATACCAAGTATATTACTGAGTTT 778
Db 61 CTTACATGGTTTCTCGATGAGAAATGTCATTTGGACATACCAAGTATATTACTGAGTTT 120

QY 779 TTTAGTCAGGGCTGGAAGCCATATGATTTACAGATCTATCTCAGGAGACACAGCAT 838
Db 121 TTTAGTCAGGGCTGGAAG- CATATGATTTACAGATCTATCTCAGGAGACACAGCAT 179

QY 839 AAGAGTTCTCGAATCGGAGAACATTATGACTGTCATCTTTGCACTACCTTCAAGAAA 898
Db 180 AAGAGTTCTCGAATCGGAGAACATTATGACTGTCATCTTTGCACTACCTTCAAGAAA 239

QY 899 CTTTGGATCAGTATTCCTCTCTAAAGTGATATAATTTTGACCTGTGTAGAACTCTCT 958
Db 240 CTTTGGATCAGTATTCCTCTCTAAAGTGATATAATTTTGACCTGTGTAGAACTCTCT 299

QY 959 GGTATACAC 967
Db 300 GGTATACAC 308

RESULT 15

US-09-976-674-4
; Sequence 4, Application US/09976674
; Patent No. 684180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674

; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117.
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 2617
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-976-674-4

Query Match 17.8%; Score 212.8; DB 3; Length 2617;
Best Local Similarity 63.9%; Pred. No. 1.4e-58;
Matches 322; Conservative 0; Mismatches 182; Indels 0; Gaps 0;

QY 404 GGGTCAATAGAAATTCACGATCAGGTGGAGACATCCCAATATCTAGCTTCTCGATATGA 463
Db 2101 GGGCAGGTGGAGATCGAGGACCGAGGTGGAGGCGCTGCAGTTCTGTGGCCGAGAAGTATGG 2160

QY 464 TTTCAATTGACTTAGATCGTGTGGGCATCCACGCTGGTCTTATGGAGGATACCTCTCCCT 523
Db 2161 CTTTCATCGACCTGAGCGGAGTTGCCATCCATGGCTGTCTACGGGGGCTTCTCTCGCT 2220

QY 524 GATGGCATTATGCGAGAGTTCAGATATCTTCAGGGTTGCTATTGCTGGGGCCCCAGTCCAC 583
Db 2221 CATGGGGCTAATCCCAAGCCCCCAGGTTTCAAGGTGGCCATCGCGGGGTGCCCCGGTCCAC 2280

QY 584 TCTGTGATCTTCTATGATACAGGATACAGGAACTTATATGGGTACCCCTGACACAGAA 643
Db 2281 CGTCTGATGGCCTTACGACACAGGTACACTGAGCGCTACATGGAGGTCCTTGAGAACAA 2340

QY 644 TGAACAGGGCTATTACTTTAGGATCTGTGGCCATGCAAGCAGAAAAAGTTCCCTCTGAACC 703
Db 2341 CCAGCAGGCTATGAGCGGGTTCCGTGGCCCTGCACGTGGAGAGCTGCCCAATGAGCC 2400

QY 704 AAATCGTTTACTGCTCTTACATGGTTTCTGGATGAGAATGCCATTTTGGCACATACCCAG 763
Db 2401 CAACCGCTTGTCTTATCTCCACGGCTTCTTGGACGAAAAACGTGCACTTTTCCACACAAA 2460

QY 764 TATATTACTGAGTTTTTTAGTGAGGCTGGAAGCCATATGATTTACAGATCTATCTCTCA 823
Db 2461 CTTCTCTGCTCCCAACTGATCCGAGCAGGAAACCTTACCAGCTCCAGATCTACCCCAA 2520

QY 824 GGAGAGACACAGCATATAGAGTTCTCTGAATCGGGAGAACATTATGAACATGCACTCTTTTGA 883
Db 2521 CGAGAGACACAGTATTCGCTGCCCCGAGTCGGCGGAGCACTATGAAGTCACTGCTCTGCA 2580

QY 884 CTACCTTCAAGAAAACTTGGATC 907
Db 2581 CTTTCTACAGGATACCTCTGAGC 2604

Search completed: May 3, 2006, 02:12:56

Job time : 388.227 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 4, 2006, 03:49:39 ; Search time 1635 Seconds
(without alignments)
2981.544 Million cell updates/sec

Title: US-10-825-632-4
Perfect score: 1197
Sequence: 1 attttgaagccacaaagac.....ttactaataaaaaaaaaa 1197

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 9306428 seqs, 2036268586 residues

Total number of hits satisfying chosen parameters: 18612856

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA New.*
1: /SID55/ptodata/2/pubpna/US08_NEW_PUB.seq1.*
2: /SID55/ptodata/2/pubpna/US06_NEW_PUB.seq.*
3: /SID55/ptodata/2/pubpna/US07_NEW_PUB.seq.*
4: /SID55/ptodata/2/pubpna/US08_NEW_PUB.seq.*
5: /SID55/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
6: /SID55/ptodata/2/pubpna/US09_NEW_PUB.seq.*
7: /SID55/ptodata/2/pubpna/US09_NEW_PUB.seq1.*
8: /SID55/ptodata/2/pubpna/US09_NEW_PUB.seq2.*
9: /SID55/ptodata/2/pubpna/US10_NEW_PUB.seq.*
10: /SID55/ptodata/2/pubpna/US10_NEW_PUB.seq1.*
11: /SID55/ptodata/2/pubpna/US10_NEW_PUB.seq2.*
12: /SID55/ptodata/2/pubpna/US10_NEW_PUB.seq3.*
13: /SID55/ptodata/2/pubpna/US10_NEW_PUB.seq4.*
14: /SID55/ptodata/2/pubpna/US11_NEW_PUB.seq.*
15: /SID55/ptodata/2/pubpna/US11_NEW_PUB.seq2.*
16: /SID55/ptodata/2/pubpna/US11_NEW_PUB.seq3.*
17: /SID55/ptodata/2/pubpna/US11_NEW_PUB.seq4.*
18: /SID55/ptodata/2/pubpna/US11_NEW_PUB.seq5.*
19: /SID55/ptodata/2/pubpna/US60_NEW_PUB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1040	86.9	3143	17	US-11-151-601-19 Sequence 19, Appl
2	778	65.0	2649	17	US-11-151-601-21 Sequence 21, Appl
C 3	118	9.9	1346	7	US-09-925-065A-669313 Sequence 669313, Appl
C 4	113.6	9.5	609	7	US-09-925-065A-743558 Sequence 743558, Appl
C 5	113.6	9.5	609	7	US-09-925-065A-743559 Sequence 743559, Appl
6	80.2	6.7	2238	18	US-11-079-463-1186 Sequence 1186, Ap
C 7	73.4	6.1	612	7	US-09-925-065A-818446 Sequence 818446, Appl
8	70	5.8	3332	18	US-11-208-288-1 Sequence 1, Appl1
9	70	5.8	3407	11	US-10-501-035-34 Sequence 34, Appl
10	68.4	5.7	2217	18	US-11-208-288-3 Sequence 3, Appl1
11	68.4	5.7	2301	10	US-10-522-789-1 Sequence 1, Appl1
12	64.4	5.4	2788	9	US-10-505-928-476 Sequence 476, Appl
13	64.4	5.4	2814	17	US-11-186-284-54 Sequence 54, Appl
14	64.4	5.4	2814	18	US-11-245-147-168 Sequence 168, Appl

15	61.8	5.2	535	18	US-11-226-869-428 Sequence 428, App
16	61.6	5.1	4852	17	US-11-136-527-2130 Sequence 2130, Ap
17	61	5.1	2283	18	US-11-208-288-5 Sequence 5, Appli
18	52.4	4.4	2778	11	US-10-932-182A-5649 Sequence 5649, Ap
19	52.4	4.4	2778	11	US-10-932-182A-5649 Sequence 5649, Ap
20	50.6	4.2	378	11	US-10-932-182A-81332 Sequence 81332, A
21	50.6	4.2	378	11	US-10-932-182A-81332 Sequence 81332, A
22	48.6	4.1	2457	11	US-10-932-182A-1107 Sequence 1107, Ap
23	48.6	4.1	2457	11	US-10-932-182A-1107 Sequence 1107, Ap
24	42	3.5	1884	18	US-11-079-463-2536 Sequence 2536, Ap
25	38	3.2	544	11	US-10-301-480-7940 Sequence 7940, Ap
26	38	3.2	544	12	US-10-301-480-7940 Sequence 7940, Ap
27	37.8	3.2	538	7	US-09-925-065A-342135 Sequence 342135, Appl
28	37.8	3.2	562	12	US-10-301-480-414923 Sequence 414923, Appl
29	37.8	3.2	562	12	US-10-301-480-1028332 Sequence 1028332, Appl
C 30	37.8	3.2	170995	17	US-11-121-086-35 Sequence 35, Appl
31	36.8	3.1	1624	10	US-10-131-826A-181 Sequence 181, App
32	36.8	3.1	1624	11	US-10-973-115B-181 Sequence 181, App
33	36.8	3.1	1624	13	US-10-137-873A-181 Sequence 181, App
34	36.8	3.1	1624	13	US-10-152-370-181 Sequence 181, App
35	36.8	3.1	1624	18	US-11-290-153-181 Sequence 181, App
C 36	36.6	3.1	487	7	US-09-925-065A-792220 Sequence 792220, Appl
C 37	36.6	3.1	487	7	US-09-925-065A-793983 Sequence 793983, Appl
C 38	36.6	3.1	487	7	US-09-925-065A-850079 Sequence 850079, Appl
C 39	36.6	3.1	538	7	US-09-925-065A-342136 Sequence 342136, Appl
40	36.6	3.1	538	7	US-09-925-065A-342137 Sequence 342137, Appl
41	36.6	3.1	562	12	US-10-301-480-414924 Sequence 414924, Appl
42	36.6	3.1	562	12	US-10-301-480-414925 Sequence 414925, Appl
43	36.6	3.1	562	12	US-10-301-480-1028333 Sequence 1028333, Appl
44	36.6	3.1	562	12	US-10-301-480-1028334 Sequence 1028334, Appl
C 45	36.6	3.1	1936	7	US-09-925-065A-83687 Sequence 83687, A

ALIGNMENTS

RESULT 1
US-11-151-601-19
; Sequence 19, Application US/11151601
; Publication No. US200600003413A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Meyers, Rachel B.
; APPLICANT: Olandt, Peter J.
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Curtis, Rory A. J.
; APPLICANT: Williamson, Mark
; APPLICANT: Welch, Nadine
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASE, PHOSPHATASE,
; AND PROTEASE FAMILY MEMBERS AND USES THEREOF
; FILE REFERENCE: MF100-054PIRCPIOMIDVIM
; CURRENT APPLICATION NUMBER: US/11/151,601
; CURRENT FILING DATE: 2005-06-13
; PRIOR APPLICATION NUMBER: US 10/170,789
; PRIOR FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: US 09/797,039
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: PCT/US01/06525
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/186,061
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 09/882,166
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: PCT/US01/19269
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/212,078
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 09/934,406
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: PCT/US01/26052
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/226,740
; PRIOR FILING DATE: 2000-08-21

```
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 3143
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (229)...(2874)
US-11-151-601-19

Query Match      86.9%; Score 1040; DB 17; Length 3143;
Best Local Similarity 89.1%; Pred. No. 1.1e-269;
Matches 1197; Conservative 0; Mismatches 0; Indels 147; Gaps 1;

QY 1 ATTTTGAAGGCACCAAGAGCTCCCTTTTAGAGCATCACCTGTAGTAGTCAGTTACGTAA 60
DB 1796 ATTTTGAAGGCACCAAGAGCTCCCTTTTAGAGCATCACCTGTAGTAGTCAGTTACGTAA 1855
QY 61 ATCTGTGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACATTTCTGTGCAATCAGTC 120
DB 1856 ATCTGTGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACATTTCTGTGCAATCAGTC 1915
QY 121 AGCACTGTGACTTCTTTATAGTAAAGTATAGTACCAAGAAATCCACACTGTGTGCC 180
DB 1916 AGCACTGTGACTTCTTTATAGTAAAGTATAGTAAACAGAGAAATCCACACTGTGTGCC 1975
QY 181 TTTACAGCTATCAAGTCTGTGAGATGACCGAAGTTCGAAACAAAGAAATTTTGGGCCA 240
DB 1976 TTTACAGCTATCAAGTCTGTGAGATGACCGAAGTTCGAAACAAAGAAATTTTGGGCCA 2035
QY 241 CCATTTTGGATTCAGCAGGTCTCTTCTGACTATACCTCTCCAGAAATTTCTCTTTTG 300
DB 2036 CCATTTTGGATTCAGCAGGTCTCTTCTGACTATACCTCTCCAGAAATTTCTCTTTTG 2095
QY 301 AAGTACTACTGGAATTAATGATGATGGAGTCTCTACAAGCCTCATGATCTACGCGTG 360
DB 2096 AAGTACTACTGGAATTAATGATGATGGAGTCTCTACAAGCCTCATGATCTACGCGTG 2155
QY 361 GAAAGAAATATCCTACTGTGCTGTTTCATATATGTTGGTCTCA----- 403
DB 2156 GAAAGAAATATCCTACTGTGCTGTTTCATATATGTTGGTCTCAAGGTGCGATGGAATA 2215
QY 404 ----- 403
DB 2216 ATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGGAATACCTAGCCTCTTAGTGTATGTGG 2275
QY 404 ----- 403
DB 2276 TTGTAGTGATAGACACAGGGGATCCTGTCCAGGAGGCTTAAATTTGAAGGCGCCTTTA 2335
QY 404 -----GGTCAAAATAGAAATTTGACGATCAGGTGGAGGACTCCAAATATCTAGCTT 453
DB 2336 AATATAAATGGGTCAAAATAGAAATTTGACGATCAGGTGGAGGACTCCAAATATCTAGCTT 2395
QY 454 CTGGATATGATTTTCATTTGACTTAGATCTGTGGGATCCAGCGTGTCTCTATGGAGGAT 513
DB 2396 CTGGATATGATTTTCATTTGACTTAGATCTGTGGGATCCAGCGTGTCTCTATGGAGGAT 2455
QY 514 ACTCTCCCTGTATGGCAATTAATGCAAGGTCAAGATATCTTCAGGGTTGCTATTGCTGGGG 573
DB 2456 ACTCTCCCTGTATGGCAATTAATGCAAGGTCAAGATATCTTCAGGGTTGCTATTGCTGGGG 2515
QY 574 CCCAGTCACTCTGTGGATCTTTCTATGATACAGGATACACGGAACGTTATATGGGTCAAC 633
DB 2516 CCCAGTCACTCTGTGGATCTTTCTATGATACAGGATACACGGAACGTTATATGGGTCAAC 2575
QY 634 CTGACCAAGATGAACAGGCTATTACTTAGGATCTGTGGCCATCGCAAGCAAGAAAGTTCC 693
DB 2576 CTGACCAAGATGAACAGGCTATTACTTAGGATCTGTGGCCATCGCAAGCAAGAAAGTTCC 2635
QY 694 CCTCTGAACCAAACTGTTTACTTGCTCTTACATGTTTCTCTGATGAGAATGTCCATTTTG 753
```

```
DB 2636 CCTCTGAACCAATCGTTTACTGCTCTTACATGGTTTCTTGATGAGAAATGTCCATTTTG 2695
QY 754 CACATACAGTATATTACTGAGTTTTTTAGTGAGGGCTGGAAGCCATATGATTTACAGA 813
DB 2696 CACATACAGTATATTACTGAGTTTTTTAGTGAGGGCTGGAAGCCATATGATTTACAGA 2755
QY 814 TCTATCCTCAGGAGAGACACAGCATTAAGAGTTTCTGAAATCGGGAGAAACATTTATGAACCTGC 873
DB 2756 TCTATCCTCAGGAGAGACACAGCATTAAGAGTTTCTGAAATCGGGAGAAACATTTATGAACCTGC 2815
QY 874 ATCTTTTGCACCTACCTTTCAAGAAAACCTTTGGATTCAGTATTGCTCTCTAAAAAGTGATAT 933
DB 2816 ATCTTTTGCACCTACCTTTCAAGAAAACCTTTGGATTCAGTATTGCTCTCTAAAAAGTGATAT 2875
QY 934 AATTTTGCACCTGTGTAGAACTCTCTGTTATACACTGGCTATTTTAACCAAAATGAGGAGGTT 993
DB 2876 AATTTTGCACCTGTGTAGAACTCTCTGTTATACACTGGCTATTTTAACCAAAATGAGGAGGTT 2935
QY 994 TAATCAACAGAAAACACAGAAATTTGATCATCATCATTTTGTATACCTGCCATGTAACATCTAC 1053
DB 2936 TAATCAACAGAAAACACAGAAATTTGATCATCATCATTTTGTATACCTGCCATGTAACATCTAC 2995
QY 1054 TCCTGAAAATTAATGTGGTGGCCATGCGAGGGTCTAGGGTTTGTGTAGTAACTTAATACC 1113
DB 2996 TCCTGAAAATTAATGTGGTGGCCATGCGAGGGTCTAGGGTTTGTGTAGTAACTTAATACC 3055
QY 1114 TTAACCCCATCCTCTCAAAATCAAAATGATACATATTTCTGAGAGACCAGCAATACCATA 1173
DB 3056 TTAACCCCATCCTCTCAAAATCAAAATGATACATATTTCTGAGAGACCAGCAATACCATA 3115
QY 1174 AGAATTAATAAAAAA----- 1197
DB 3116 AGAATTAATAAAAAA----- 3139

RESULT 2
US-11-151-601-21
; Sequence 21, Application US/11151601
; Publication No. US2006003413A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Olandt, Peter J.
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Curtis, Rory A. J.
; APPLICANT: Williamson, Mark
; APPLICANT: Welch, Nadine
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASE, PHOSPHATASE,
; FILE REFERENCE: AND PROTEASE FAMILY MEMBERS AND USES THEREOF
; CURRENT APPLICATION NUMBER: US/11/151,601
; PRIOR FILING DATE: 2005-06-13
; PRIOR APPLICATION NUMBER: US 10/170,789
; PRIOR FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: US 09/797,039
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: PCT/US01/06525
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/186,061
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 09/882,166
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: PCT/US01/19269
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/212,078
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 09/934,406
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: PCT/US01/26052
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/226,740
; PRIOR FILING DATE: 2000-08-21
```

```
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 2649
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-151-601-21

Query Match      65.0%; Score 778; DB 17; Length 2649;
Best Local Similarity 86.4%; Pred. No. 4.7e-199;
Matches 935; Conservative 0; Mismatches 0; Indels 147; Gaps 1;

QY 1 ATTTGAAGGACCAAGAGACTCCCTTTAGAGCATCACCTGTAGTACAGTACGTTAA 60
Db 1568 ATTTTGAAGGACCAAGAGACTCCCTTTAGAGCATCACCTGTAGTACAGTACGTTAA 1627

QY 61 ATCCTGGAGAGGTGACAGGCTGACTGACCGTGGCTACTACATTCCTGCTGCATCAGTC 120
Db 1628 ATCCTGGAGAGGTGACAGGCTGACTGACCGTGGCTACTACATTCCTGCTGCATCAGTC 1687

QY 121 AGCAGTGTGACTCTTTATAGTAAGTATAGTAACAGAGAAATCCACACTGTGTGCC 180
Db 1688 AGCAGTGTGACTCTTTATAGTAAGTATAGTAACAGAGAAATCCACACTGTGTGCC 1747

QY 181 TTTACAAGCTATCAAGTCTGAAGATGACCCAACTTGCAAAACAAAGGAATTTTGGGCA 240
Db 1748 TTTACAAGCTATCAAGTCTGAAGATGACCCAACTTGCAAAACAAAGGAATTTTGGGCA 1807

QY 241 CCATTTTGGATTGACGAGGTCTCTTCTGACTATATCTCTCAGAAATTTTCTTTTG 300
Db 1808 CCATTTTGGATTGACGAGGTCTCTTCTGACTATATCTCTCAGAAATTTTCTTTTG 1867

QY 301 AAGTACTACTGGAATTAATGTATGGATGCTCTACAGCCTCATGATCTACAGCCTG 360
Db 1868 AAGTACTACTGGAATTAATGTATGGATGCTCTACAGCCTCATGATCTACAGCCTG 1927

QY 361 GAAAGAAATATCCCTACTGCTGCTTATATATATGTTGCTCTCA- 403
Db 1928 GAAAGAAATATCCCTACTGCTGCTTATATATATGTTGCTCTCAAGTGGCAGTTGGAATA 1987

QY 404 ----- 403
Db 1988 ATCGGTTTAAAGGAGTCAAGTATTTTCGCTTGAATACCCCTAGCCTCTCTAGGTATGTGG 2047

QY 404 ----- 403
Db 2048 TTGTAGTGTAGACACAGGGGATCCTGTACCGAGGGCTTAAATTTGAAGGGCTTTA 2107

QY 404 -----GGGTCAAAATAGAAATTCAGCATCAGGTGGAGACTCCAAATATCTAGCTT 453
Db 2108 AATATAAAATGGGTCAAAATAGAAATTCAGCATCAGGTGGAGACTCCAAATATCTAGCTT 2167

QY 454 CTGATATGATTTTCAATGACTATAGATCGTGTGGGCATCCAGCGCTGGTCTATGAGGAT 513
Db 2168 CTGATATGATTTTCAATGACTATAGATCGTGTGGGCATCCAGCGCTGGTCTATGAGGAT 2227

QY 514 ACCTCTCCCTGATGCTAATGACAGGTCAGATATCTTCAGGTTGCTATTGCTGGGG 573
Db 2228 ACCTCTCCCTGATGCTAATGACAGGTCAGATATCTTCAGGTTGCTATTGCTGGGG 2287

QY 574 CCCAGTCACTCTGTGGATCTTCTATGATACAGGATACCGGAACGTTATATGGGTACCC 633
Db 2288 CCCAGTCACTCTGTGGATCTTCTATGATACAGGATACCGGAACGTTATATGGGTACCC 2347

QY 634 CTGACCAAGTAAACAGGGCTATTACTAGGATCTGTGGCCATGCAAGCAGAAAAAGTTCC 693
Db 2348 CTGACCAAGTAAACAGGGCTATTACTAGGATCTGTGGCCATGCAAGCAGAAAAAGTTCC 2407

QY 694 CCTCTGAACCAATCGTTTACTGTCTTACATGTTTCTTGGATGAGATGTCATTG 753
Db 2408 CCTCTGAACCAATCGTTTACTGTCTTACATGTTTCTTGGATGAGATGTCATTG 2467
```

```
QY 754 CACATACCAGTATATTAAGTGGTGTGAGGGCTGAAAGCCATATGATTTACAGA 813
Db 2468 CACATACCAGTATATTAAGTGGTGTGAGGGCTGAAAGCCATATGATTTACAGA 2527

QY 814 TCTATCCTCAGGAGACACAGCATATAGAGTTCTGTAATCGGAGAACATTATGAACCTGC 873
Db 2528 TCTATCCTCAGGAGACACAGCATATAGAGTTCTGTAATCGGAGAACATTATGAACCTGC 2587

QY 874 ATCTTTTGCACCTACTCTTCAAGAAAACCTTGGATCAGCTATTTGCTCTTAAAGTGTATAT 933
Db 2588 ATCTTTTGCACCTACTCTTCAAGAAAACCTTGGATCAGCTATTTGCTCTTAAAGTGTATAT 2647

QY 934 AA 935
Db 2648 AA 2649

RESULT 3
US-09-925-065A-669313/c
; Sequence 669313, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 669313
; LENGTH: 1346
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-669313

Query Match      9.9%; Score 118; DB 7; Length 1346;
Best Local Similarity 100.0%; Pred. No. 3.2e-21;
Matches 118; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 256 CAGGTCTCTTCTCCTGACTATATCTCTCCAGAAATTTTCTCTTTTGAAGTACTCTGGAT 315
Db 118 CAGGTCTCTTCTCCTGACTATATCTCTCCAGAAATTTTCTCTTTTGAAGTACTCTGGAT 59

QY 316 TTACATTTGATGGATGCTCTACAGCCTCATGATCTACAGCCTGGAAGAAATATCC 373
Db 58 TTACATTTGATGGATGCTCTACAGCCTCATGATCTACAGCCTGGAAGAAATATCC 1

RESULT 4
US-09-925-065A-743558/c
; Sequence 743558, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
```

```
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 743558
; LENGTH: 609
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-743558

Query Match          9.5%; Score 113.6; DB 7; Length 609;
Best Local Similarity 96.7%; Pred. No. 3.6e-20;
Matches 116; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 695 CTCGAACCAATCGTTTACTGCTCTTACATGGTTTCTCGGATGAGAATGTCATTTCG 754
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 205 CTCGAGACCAATCGTTTACTGCTCTTACATGGTTTCTCGGATGAGAATGTCATTTCG 146

Qy 755 ACATACCAGTATATTACTGAGTGTCTTTAGTGAGGCTGGAAAGCCCATATGATTTACAGAT 814
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 145 ACATACCAGTATATTACTGAGTGTCTTTAGTGAGGCTGGAAAGCCCATATGATTTACAGT 86

RESULT 5
US-09-925-065A-743559/c
; Sequence 743559, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 10827,135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 743559
; LENGTH: 609
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-743559

Query Match          9.5%; Score 113.6; DB 7; Length 609;
Best Local Similarity 96.7%; Pred. No. 3.6e-20;
Matches 116; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 695 CTCGAACCAATCGTTTACTGCTCTTACATGGTTTCTCGGATGAGAATGTCATTTCG 754
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 205 CTCGAGACCAATCGTTTACTGCTCTTACATGGTTTCTCGGATGAGAATGTCATTTCG 146

Qy 755 ACATACCAGTATATTACTGAGTGTCTTTAGTGAGGCTGGAAAGCCCATATGATTTACAGAT 814
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 145 ACATACCAGTATATTACTGAGTGTCTTTAGTGAGGCTGGAAAGCCCATATGATTTACAGT 86

RESULT 6
US-11-079-463-1186
; Sequence 1186, Application US/11079463
```

```
; Publication No. US20060073161A1
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTEROIDES FRAC
; FILE REFERENCE: FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/11/079,463
; CURRENT FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/128,705
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: US 09/540,209
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 10444
; SEQ ID NO 1186
; LENGTH: 2238
; TYPE: DNA
; ORGANISM: B.fragilis
US-11-079-463-1186

Query Match          6.7%; Score 80.2; DB 18; Length 2238;
Best Local Similarity 50.7%; Pred. No. 6e-11;
Matches 252; Conservative 0; Mismatches 233; Indels 12; Gaps 2;

Qy 402 CAGGGTCAAAATAGAATTGACGATCAGTGGAAGGACTCCAAATATCTAGCTTCTCGATAT 461
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1747 CTGGGTGTAAGAAGACTAAAGACGAGTGGAAGCTGCCAAATATCTGG---GTGACTG 1803

Qy 462 GATTTTCATGACTTAGATCGTGGGCATCCACGGCTGGTCTTATGAGGATACCTCTCC 521
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1804 CCTATGTGGACAAAGGACGATTGGTATCTGGGGATGGAGTTTCGGCGGATATATGACC 1863

Qy 522 CTGATGGCATTAATGACAGAGTCAGATATCTTCAGGGTCTGCTATTTGTTGGGGCCCGAGTC 581
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1864 ATCATGAGTATGAGCGAAGGTACACCCGTGTTAAAGCCGGAGTTGCTGTGGCCGACCT 1923

Qy 582 ACTCTGTGGATCTTCTATGATACAGGATACACGGAACGTTTATATGGTTCACCTGCACAG 641
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1924 ACAGACTGGAATATATACGATACAGATATATACCGAACGCTTTATGCGCAGCCGCAAGAA 1983

Qy 642 AATGAACAGGCTTATCTTAGGATCTGTGGCCATGCAAGCAGAAAGTTTCCCTCTGAA 701
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1984 AATGCCGAAGGCTATAAAGCAGCTTCAGCATTCAGCCGTGCAGAACCTGCATGTTAAC 2043

Qy 702 CCAAAATCGTTTACTGCTCTTACATGGTTTCTCGATGAGAATGTCATTTTGCACATACC 761
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2044 C-----TGCTCTTTGTACACGGTATGCGCAGATGATAATGTTTCACTTCCAGAACTGT 2094

Qy 762 AGTATATTACTGAGTGTCTTTAGTGAGGCTGGAAAGCCATATGATTTTACAGATCTATCCT 821
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2095 ACAGAAATATGACAGACACCTGTTACAACTCGGAAACAGTTCGATATGAGGTATACACC 2154

Qy 822 CAGGAGACACACATGAAGAGTTCTTGAATCGGAGAACATTATGAACCTGCATCTTTTG 881
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2155 AACCAGAAATCATAGCATCTATGTTGGAATAACCCGTAACCCATTGTATACGAAGCTGACG 2214

Qy 882 CACTACCTTCAAGAAA 898
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2215 AACTTCTTCCGGAATAA 2231
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 7
US-09-925-065A-818446/c
; Sequence 818446, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 10827,135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
```

;; PRIOR APPLICATION NUMBER: US 60/252,147
;; PRIOR FILING DATE: 2000-11-20
;; PRIOR APPLICATION NUMBER: US 60/250,092
;; PRIOR FILING DATE: 2000-11-30
;; PRIOR APPLICATION NUMBER: US 60/261,766
;; PRIOR FILING DATE: 2001-01-16
;; PRIOR APPLICATION NUMBER: US 60/289,846
;; PRIOR FILING DATE: 2001-05-09
;; NUMBER OF SEQ ID NOS: 957086
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 818446
;; LENGTH: 612
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-09-925-065A-818446

Query Match 6.1%; Score 73.4; DB 7; Length 612;
Best Local Similarity 98.7%; Pred. No. 2.5e-09;
Matches 74; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 740 GAATGTCATTTCACATACCAGTATATTAATGAGGCTGGAAAGCC 799
Db |||||||||||||||||||||||||||||||||||||||||||||||||||
612 GAATGTCATTTCACATACCAGTATATTAATGAGGCTGGAAAGCC 553
QY 800 ATATGATTACAGAT 814
Db |||||||||||||||
552 ATATGATTACAGGT 538

RESULT 8
US-11-208-288-1
;; Sequence 1, Application US/11208288
;; Publication No. US20060051366A1
;; GENERAL INFORMATION:
;; APPLICANT: CHANG, Chiwen
;; TITLE OF INVENTION: USE OF SOLUBLE CD26 AS INHIBITOR OF
;; FILE REFERENCE: 39533-0001
;; CURRENT APPLICATION NUMBER: US/11/208,288
;; CURRENT FILING DATE: 2005-08-18
;; PRIOR APPLICATION NUMBER: US 60/605,013
;; PRIOR FILING DATE: 2004-08-26
;; NUMBER OF SEQ ID NOS: 8
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 1
;; LENGTH: 3332
;; TYPE: DNA
;; ORGANISM: homo sapiens
US-11-208-288-1

Query Match 5.8%; Score 70; DB 18; Length 3332;
Best Local Similarity 47.7%; Pred. No. 3.9e-08;
Matches 268; Conservative 0; Mismatches 290; Indels 4; Gaps 2;
QY 402 CAGGTCATAATAGAAATTGACATCAGGTGGGAAGGACTCCATATCTAGCTTCTCGATAT 461
Db 1792 CTGGGAACATTTGAAGTTGAAGATCAAAATTGAAGCAGCCAGA---CAATTTTCAAAAATG 1848
QY 462 GATTTTCATTGACTTAGATCGTGGGCATCCACGGCTGCTCTATGGAGGATACCTCTCC 521
Db 1849 GGAATTTGGGACAAACAAACGAATTTGCAATTTGGGGCTGGTCATATGAGGGTACGTAACC 1908
QY 522 CTGATGGCATTAATGACAGAGTCCAGATATCTTCAGGGTTGCTATTGCTGGGGCCCCCAGTC 581
Db 1909 TCAATGGTCTGGGATCGGGAAGTGGCGTGTTCAGTGTGGATAGCCGTGGCGCTGTA 1968
QY 582 ACTCTGTGGATCTTTCTATGATACAGGATACACGGAAAGTATATGGGTACCCCTGACCCAG 641
Db 1969 TCCCGGTGGGAGTACTATGACTCAGTGTACACAGAAGTTATAGGGTCTCCCAACTCCA 2028
QY 642 AATGAACAGGGCTATTACTTAGGATCTGTGGCCATCAAGCAGAGAAAAGTTCCCTCTGAA 701
Db 2029 GAAGCAACCTTGAACCATTAAGAAATTCACAGTCTATGAGCAGAGCTGAAAATTTTAA 2088

QY 702 CCAAAATCGTTTACTGCTCTTACATGGTTCTCTGATGAGAATGTCAATTTGCATATACC 761
Db 2089 CAAAGTTGAGTACCTCTTATTCATGGAAACAGCAGATGATAACGTTCACTTTTCAGCAGTCA 2148
QY 762 AGTATATTACTGAGTCTTTTGTAGTCAGGGCTGGAAAGCCATATGATTTACAGATCTATCCT 821
Db 2149 GCTCAGATCTCCAAAGCCCTGGTCGATGTTGGAGTGGATTTCCAGGCAATGGGTATCT 2208
QY 822 CAGGAGACACAGACATAGAAGTTCTCTGAATCGGGAAGAACATTTATGAATGCAATCTTTTG 881
Db 2209 GATGAGACCATGGAATAGCTAGCAGACAGCAGACACATATATATACCCACATGAGC 2268
QY 882 CACTACCTTCAAGAAAACCTTTGGATCAGGTATTTGCTGCTCTAAAAGTATATAATTTGA 941
Db 2269 CACTTCAATAAAACAATGTTCTCTTTACCT-TAGCACCTCAAAATACCATGCAATTTAAA 2327
QY 942 CCTGTGTAGAACTCTCTGGTAT 963
Db 2328 GCTTATTAAACTCATTTTGT 2349

RESULT 9
US-10-501-035-34
;; Sequence 34, Application US/10501035
;; Publication No. US20060046249A1
;; GENERAL INFORMATION:
;; APPLICANT: Bristol-Myers Squibb Company
;; TITLE OF INVENTION: IDENTIFICATION OF POLYNUCLEOTIDES AND POLYPEPTIDE FOR PREDICTING
;; TITLE OF INVENTION: ACTIVITY OF COMPOUNDS THAT INTERACT WITH PROTEIN TYROSINE KINASE
;; TITLE OF INVENTION: AND/OR PROTEIN TYROSINE KINASE PATHWAYS
;; FILE REFERENCE: D0185 PCT
;; CURRENT APPLICATION NUMBER: US/10/501,035
;; CURRENT FILING DATE: 2004-07-09
;; PRIOR APPLICATION NUMBER: US 60/350,061
;; PRIOR FILING DATE: 2002-01-18
;; NUMBER OF SEQ ID NOS: 795
;; SOFTWARE: PatentIn version 3.2
;; SEQ ID NO 34
;; LENGTH: 3407
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-10-501-035-34
Query Match 5.8%; Score 70; DB 11; Length 3407;
Best Local Similarity 47.7%; Pred. No. 4e-08;
Matches 268; Conservative 0; Mismatches 290; Indels 4; Gaps 2;
QY 402 CAGGTCATAATAGAAATTGACATCAGGTGGGAAGGACTCCATATCTAGCTTCTCGATAT 461
Db 1867 CTGGGAACATTTGAAGTTGAAGATCAAAATTGAAGCAGCCAGA---CAATTTTCAAAAATG 1923
QY 462 GATTTTCATTGACTTAGATCGTGGGCATCCACGGCTGCTCTATGGAGGATACCTCTCC 521
Db 1924 GGAATTTGGACAAACAAACGAATTTGCAATTTGGGGCTGGTCATATGAGGGTACGTAACC 1983
QY 522 CTGATGGCATTAATGACAGAGTCCAGATATCTTCAGGGTTGCTATTGCTGGGGCCCCCAGTC 581
Db 1984 TCAATGGTCTGGGATCGGGAAGTGGCGTGTTCAGTGTGGATAGCCGTGGCGCTGTA 2043
QY 582 ACTCTGTGGATCTTTCTATGATACAGGATACACGGAAAGTATATGGGTACCCCTGACCCAG 641
Db 2044 TCCCGGTGGGAGTACTATGACTCAGTGTACACAGAAGTTTACATGGGTCTCCCAACTCCA 2103
QY 642 AATGAACAGGGCTATTACTTAGGATCTGTGGCCATCAAGCAGAGAAAAGTTCCCTCTGAA 701
Db 2104 GAAGCAACCTTGAACCATTAAGAAATTCACAGTCTATGAGCAGAGCTGAAAATTTTAA 2163
QY 702 CCAAAATCGTTTACTGCTCTTACATGGTTCTCTGATGAGAATGTCCATTTTGCATATACC 761
Db 2164 CAAAGTTGAGTACCTCTCTTTATTCATGGNAACAGCAGATGATAAGTTCACTTTTCAGCAGTCA 2223
QY 762 AGTATATTACTGAGTCTTTTGTAGTGGGCTGGAAAGCCATATGATTTACAGATCTATCCT 821

Db 2224 GCTCAGATCTCCAAAGCCCTGGTGGATGTTGGAGTGGATTTCCAGGCAATGTGGTATCT 2283
Qy 822 CAGGAGAGACACAGCATAAGAGTCTCTGAATCGGAGAACATTTATGAACCTGCATCTTTTG 881
Db 2284 GATGAAGACCATGGAATAGCTAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 2343
Qy 882 CACTTACCTTCAAGAAACCTTGGATCAGCTATTTGCTCTTAAAGTGATATAATTTTGA 941
Db 2344 CACTTCATAAACAATGTTCTCTTTACCT-TAGCACCTCAAAATACCATGCCATTAA 2402
Qy 942 CCTGTGTAGAACTCTCTGGTAT 963
Db 2403 GCTATTAAACACTATTTTGT 2424

RESULT 10

US-11-208-288-3

; Sequence 3, Application US/11208288

; Publication No. US20060031366A1

; GENERAL INFORMATION:

; APPLICANT: CHANG, Chiwen

; TITLE OF INVENTION: USE OF SOLUBLE CD26 AS INHIBITOR OF

; FILE REFERENCE: 39533-0001

; CURRENT APPLICATION NUMBER: US/11/208,288

; CURRENT FILING DATE: 2005-08-18

; PRIOR APPLICATION NUMBER: US 60/605,013

; PRIOR FILING DATE: 2004-08-26

; NUMBER OF SEQ ID NOS: 8

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 3

; LENGTH: 2217

; TYPE: DNA

; ORGANISM: Homo sapiens

US-11-208-288-3

Query Match 5.7%; Score 68.4; DB 18; Length 2217;

Best Local Similarity 48.9%; Pred. No. 9.1e-08;

Matches 214; Conservative 0; Mismatches 221; Indels 3; Gaps 1;

Qy 402 CAGGTCAAATAGAAATTGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGCATAT 461
Db 1708 CTGGGAACATTTGAAGTTGAAGATCAAAATTGAAGCAGCCAGA---CAATTTTCAAAAATG 1764
Qy 462 GATTTCATTCAGTTAGATCTGTGGGCATCCAGCGCTGCTCTATGGAGGATACCTCTCC 521
Db 1765 GGAATTTGGACAAACAACGAATTGCAATTTGGGGCTGGTCATATGGAGGTACGTAAAC 1824
Qy 522 CTGATGGCATTAATGCAGAGTCAGATATCTTCAGGGTTGCTATTGTGGGGCCCAAGTC 581
Db 1825 TCAATGGTCTGGGATCGGAAGTGGCGGTTCNAAGTGGNATAGCGTGGCGCCCTGTA 1884
Qy 582 ACTCTGTGGATCTTCTATGATACAGGATACAGGAACGTTATATGGGTCAACCTTGACAG 641
Db 1885 TCCCGGTGGGAGTACTATGACTCAGGTGATACAGAACGTTTACATGGGTTCTCCCAACTCCA 1944
Qy 642 AATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTGAA 701
Db 1945 GAAGACAACCTTGACCAATTACAGAAAATTCAAAGTCATGAGCAGAGCTGAAAAATTTTAA 2004
Qy 702 CCAATCGTTTACTGCTCTTACATGGTTTCTGGATGAGATGTCATTTTGGCATAVACC 761
Db 2005 CAAGTTAGTACCTCTCTTATTCATGGAACAGCAGATGATAAGCTTCACTTTTCCAGAGTCA 2064
Qy 762 AGTATATTACTAGTATTTTATGTCAGGGCTGGAAAGCCATATGATTTTACAGATCTATCCT 821
Db 2065 GCTCAGATCTCCAAAGCCCTGGTGGATGTTGGAGTGGATTTCCAGGCAATGTGGTATCT 2124
Qy 822 CAGGAGAGACACAGCATA 839
Db 2125 GATGAAGACCATGGAATA 2142

RESULT 11

US-10-522-789-1

; Sequence 1, Application US/10522789

; Publication No. US20050260732A1

; GENERAL INFORMATION:

; APPLICANT: TANABE SEIYAKU CO., LTD.

; TITLE OF INVENTION: Three-dimensional structure of dipeptidyl peptidase IV

; FILE REFERENCE: 03-039-PCT

; CURRENT APPLICATION NUMBER: US/10/522,789

; CURRENT FILING DATE: 2005-01-28

; PRIOR APPLICATION NUMBER: US 60/398,761

; PRIOR FILING DATE: 2002-07-29

; NUMBER OF SEQ ID NOS: 2

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 1

; LENGTH: 2301

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (1)..(2301)

; OTHER INFORMATION:

US-10-522-789-1

Query Match 5.7%; Score 68.4; DB 10; Length 2301;

Best Local Similarity 48.9%; Pred. No. 9.2e-08;

Matches 214; Conservative 0; Mismatches 221; Indels 3; Gaps 1;

Qy 402 CAGGTCAAATAGAAATTGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGCATAT 461
Db 1792 CTGGGAACATTTGAAGTTGAAGATCAAAATTGAAGCAGCCAGA---CAATTTTCAAAAATG 1848
Qy 462 GATTTCATTCAGTTAGATCTGTGGGCATCCAGCGCTGCTCTATGGAGGATACCTCTCC 521
Db 1849 GGAATTTGGACAAACAACGAATTTGCAATTTGGGGCTGGTCATATGGAGGTACGTAAAC 1908
Qy 522 CTGATGGCATTAATGCAGAGGTCAAGATATCTTCAGGGTTGCTATTGTGGGGCCCAAGTC 581
Db 1909 TCAATGGTCTCGGGATCGGGAAGTGGCGTGTTCAGATGTGGAATAGCGTGGCGCCCTGTA 1968
Qy 582 ACTCTGTGGATCTTCTATGATACAGGATACAGGAACGTTATATGGGTCAACCTTGACAG 641
Db 1969 TCCCGGTGGGAGTACTATGACTCAGGTGATACAGAACGTTTACATGGGTCTCCCAACTCCA 2028
Qy 642 AATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTGAA 701
Db 2029 GAAGACAACCTTGACCAATTACAGAAATTCACAGTCATGAGCAGAGCTGAAAAATTTTAA 2088
Qy 702 CCAATCGTTTACTGCTCTTACATGGTTTCTGGATGAGAAATGTCCATTTTGCACATACC 761
Db 2089 CAAGTTAGTACCTCTCTTATTCATGGAACAGCAGATGATAACGTTTCACTTTTCCAGAGTCA 2148
Qy 762 AGTATATTACTAGTATTTTATGTCAGGGCTGGAAAGCCATATGATTTTACAGATCTATCCT 821
Db 2149 GCTCAGATCTCCAAAGCCCTGGTGGATGTTGGAGTGGATTTCCAGGCAATGTGGTATCT 2208
Qy 822 CAGGAGAGACACAGCATA 839
Db 2209 GATGAAGACCATGGAATA 2226

RESULT 12

US-10-505-928-476

; Sequence 476, Application US/10505928

; Publication No. US20060088532A1

; GENERAL INFORMATION:

; APPLICANT: Ludwig Institute for Cancer Research et al.

; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES

; FILE REFERENCE: 28967/39178

; CURRENT APPLICATION NUMBER: US/10/505,928

; CURRENT FILING DATE: 2004-08-27

; PRIOR APPLICATION NUMBER: US 60/363,019

```
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: PatentIn 3.2
; SEQ ID NO 476
; LENGTH: 2788
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-505-928-476

Query Match      5.4%; Score 64.4; DB 9; Length 2788;
Best Local Similarity 48.1%; Pred. No. 1.2e-06;
Matches 182; Conservative 0; Mismatches 196; Indels 0; Gaps 0;

QY 462 GATTTCATTGACTTAGATCGTGTGGGCATCCACGGCTGCTCTATGAGGATACCTCTCC 521
Db 2039 GGTTTCATTGATGAAAAAAGATAGCCATATGGGGCTGGTCTATGAGGATACGTTTCA 2098

QY 522 CTGATGCAATTAATGCAGAGGTGAGATATCTTCAGGGTTGCTATTGCTGGGGCCCCAGTC 581
Db 2099 TCACCTGGCCCTTGCATCTGGAATCGTCTTTTCAAATGATGATGAGTGGCTCCAGTC 2158

QY 582 ACTCTGTGGATCTTCTATGATACAGGATACAGGAAACGTTATATGGGTCACCCCTGACAG 641
Db 2159 TCCAGCTGGGAATATTACGGCTCTGTCTACACAGAGATTCATGGGTCTCCCAACAAG 2218

QY 642 AATGAAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAAGAAAGTTCCCTCTGAA 701
Db 2219 GATGATAATCTTGAGCACTATAAGAATTCAACTGTGATGGCAAGACAGAGATATTTTCA 2278

QY 702 CCAATCGTTTACTGCTCTTACATGGTTCCTGGATGAGAATGTCCATTTTGACATACC 761
Db 2279 AATGTAGACTATCTTCTCATCCAGCAACAGCAGATGATAATGTGCACCTTTTCAAACTCA 2338

QY 762 AGTATATTACTAGATTTTTTTAGTGAGGGCTGGAAAGCCATATGATTTACAGATCTATCCT 821
Db 2339 GCACAGATTGCTAAAGCTCTGTTAATGCAAGTGGATTTCCAGGCAATGTGTACTCT 2398

QY 822 CAGGAGACACAGCATATA 839
Db 2399 GACCAGAACCCAGGCTTA 2416

RESULT 13
US-11-186-284-54
; Sequence 54, Application US/11186284
; Publication No. US20050266493A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Berger, Allison
; APPLICANT: Guillemette, Tracy L.
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Thibodeau, Stephen N.
; APPLICANT: Burgart, Lawrence J.
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF COLON CANCER
; FILE REFERENCE: MM01-029P2ENM
; CURRENT APPLICATION NUMBER: US/11/186,284
; CURRENT FILING DATE: 2005-07-21
; PRIOR APPLICATION NUMBER: US/10/301,822
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 60/339,971
; PRIOR FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: US 60/361,978
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/381,988
; PRIOR FILING DATE: 2002-05-20
; NUMBER OF SEQ ID NOS: 228
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 54
; LENGTH: 2814
```

```
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (209)...(2491)
US-11-186-284-54

Query Match      5.4%; Score 64.4; DB 17; Length 2814;
Best Local Similarity 48.1%; Pred. No. 1.2e-06;
Matches 182; Conservative 0; Mismatches 196; Indels 0; Gaps 0;

QY 462 GATTTCATTGACTTAGATCGTGTGGGCATCCACGGCTGCTCTATGAGGATACCTCTCC 521
Db 2039 GGTTTCATTGATGAAAAAAGATAGCCATATGGGGCTGGTCTATGAGGATACGTTTCA 2098

QY 522 CTGATGCAATTAATGCAGAGGTGAGATATCTTCAGGGTTGCTATTGCTGGGGCCCCAGTC 581
Db 2099 TCACCTGGCCCTTGCATCTGGAATCGTCTTTTCAAATGATGATGAGTGGCTCCAGTC 2158

QY 582 ACTCTGTGGATCTTCTATGATACAGGATACAGGAAACGTTATATGGGTCACCCCTGACAG 641
Db 2159 TCCAGCTGGGAATATTACGGCTCTGTCTACACAGAGATTCATGGGTCTCCCAACAAG 2218

QY 642 AATGAAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAAGAAAGTTCCCTCTGAA 701
Db 2219 GATGATAATCTTGAGCACTATAAGAATTCAACTGTGATGGCAAGACAGAGATATTTTCA 2278

QY 702 CCAATCGTTTACTGCTCTTACATGGTTCCTGGATGAGAATGTCCATTTTGACATACC 761
Db 2279 AATGTAGACTATCTTCTCATCCAGCAACAGCAGATGATAATGTGCACCTTTTCAAACTCA 2338

QY 762 AGTATATTACTAGATTTTTTTAGTGAGGGCTGGAAAGCCATATGATTTACAGATCTATCCT 821
Db 2339 GCACAGATTGCTAAAGCTCTGTTAATGCAAGTGGATTTCCAGGCAATGTGTACTCT 2398

QY 822 CAGGAGACACAGCATATA 839
Db 2399 GACCAGAACCCAGGCTTA 2416

RESULT 14
US-11-245-147-168
; Sequence 168, Application US/11245147
; Publication No. US2006030541A1
; GENERAL INFORMATION:
; APPLICANT: GARCIA, TERESA
; APPLICANT: ROMAN ROMAN, SERGIO
; APPLICANT: BARON, ROLAND
; APPLICANT: CALL, KATHERINE
; APPLICANT: THEILHABER, JOACHIM
; APPLICANT: CONNOLLY, TIMOTHY
; APPLICANT: JACKSON, AMANDA
; APPLICANT: BUSHNELL, STEVEN
; APPLICANT: RAWADI, GEORGES
; TITLE OF INVENTION: GENES INVOLVED IN OSTEOGENESIS, AND METHODS OF USE
; FILE REFERENCE: 37991-0023
; CURRENT APPLICATION NUMBER: US/11/245,147
; CURRENT FILING DATE: 2005-10-07
; PRIOR APPLICATION NUMBER: PCT/IB02/02211
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: 60/281,400
; PRIOR FILING DATE: 2001-04-05
; NUMBER OF SEQ ID NOS: 246
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 168
; LENGTH: 2814
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Homo sapiens fibroblast activation protein, alpha
; OTHER INFORMATION: (FAP), mRNA
US-11-245-147-168
```

```
Query Match      5.4%; Score 64.4; DB 18; Length 2814;
Best Local Similarity 48.1%; Pred. No. 1.2e-06;
Matches 182; Conservative 0; Mismatches 196; Indels 0; Gaps 0;

Qy 462 GATTTCATTGACCTTAGATCGTGTGGGCATCCACGGCTGGTCTCTATGGAGGATACCTCTCC 521
Db 2039 GGTTCATTGATGAAAAAGAAATAGCCATATGGGGCTGGTCTCTATGGAGGATACGTTTCA 2098

Qy 522 CTGATGGCATTAAATGACAGAGGTGAGATATCTTTCAGGGTTGCTATTGTCTGGGGCCCCAGTC 581
Db 2099 TCACCTGGCCCTTGCACTGTGAACTGGTCTTTTCAAATGGGTATAGCAGTGGCTCCAGTC 2158

Qy 582 ACTCTGTGGATCTTCTATGATACAGGATACAGGAACGTATATATGGGTCAACCTGACCAG 641
Db 2159 TCCAGCTGGGAATATTACGGCTCTGTCTACACAGAGAGATTTCATGGGTCTCCCAACAAAG 2218

Qy 642 AATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAGAAAGTTCCCTCTGAA 701
Db 2219 GATGATAATCTTGAGCACTATAAGAAATTCAACTGTGATGGCAAGAGCAGAATATTTCCAG 2278

Qy 702 CCAATCGTTTACTGCTCTTACATGGTTCCTTGGATGAGAAATGTCATTTTGCACATACC 761
Db 2279 AATGTAGACTATCTTCTCATCCAGGAACAGCAGATGATATGTGCACCTTCAAACTCA 2338

Qy 762 AGTATATTACTGAGTTTATTAGTGGGCTGGAAAGCCATATGATTTACAGATCTATCTT 821
Db 2339 GCACAGATTGCTAAGCTCTGGTTAATGCAAGTGGATTTCCAGGCATATGGTACTCT 2398

Qy 822 CAGGAGACACACAGCAT 839
Db 2399 GACCAGAACCCAGCGCTTA 2416

RESULT 15
US-11-226-869-428
; Sequence 428, Application US/11226869
; Publication No. US2006069054A1
; GENERAL INFORMATION:
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Sleath, Paul R.
; APPLICANT: Persing, David H.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C14
; CURRENT APPLICATION NUMBER: US/11/226,869
; CURRENT FILING DATE: 2005-09-13
; NUMBER OF SEQ ID NOS: 627
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 428
; LENGTH: 535
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-226-869-428

Query Match      5.2%; Score 61.8; DB 18; Length 535;
Best Local Similarity 58.4%; Pred. No. 3.2e-06;
Matches 108; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

Qy 462 GATTTCATTGACCTTAGATCGTGTGGGCATCCACGGCTGGTCTCTATGGAGGATACCTCTCC 521
Db 273 GGTTCATTGATGAAAAAGAAATAGCCATATGGGGCTGGTCTCTATGGAGGATACGTTTCA 332

Qy 522 CTGATGGCATTAAATGACAGAGGTGAGATATCTTTCAGGGTTGCTATTGTCTGGGGCCCCAGTC 581
Db 333 TCACCTGGCCCTTGCACTGTGAACTGGTCTTTTCAAATGGGTATAGCAGTGGCTCCAGTC 392

Qy 582 ACTCTGTGGATCTTCTATGATACAGGATACAGGAACGTATATATGGGTCAACCTGACCAG 641
Db 393 TCCAGCTGGGAATATTACGGCTCTGTCTACACAGAGAGATTTCATGGGTCTCCCAACAAAG 452

Qy 642 AATGA 646
Db 453 GATGA 457
```

Search completed: May 4, 2006, 09:26:07
Job time : 1637 secs

GenCore version 5.1.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: May 2, 2006, 01:03:49 ; Search time 8.12347 Seconds
(without alignments)
2835.522 Million cell updates/sec

Title: US-10-825-632-4
Perfect score: 2140
Sequence: 1 attttgaaggccaaaagac.....ttactaaaaaaaaaaaaa 1197

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0
Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 566832

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ n2p.model -DEV=x1p
-O=/abs/ABSSWEB spool/US10825632/runat_01052006.105948_3262/app_query_fasta_1
-DB=PIR -QFMT=fastan -SUFFIX=n2p.rpr -MINMATCH=0.1 -LOOPECL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs02p
-USER=US10825632 @CGN 1.1.92 @runat_01052006.105948_3262 -NCPU=6 -ICPU=3
-NO_MMAP -NEG SCORES=0 -WAIT -DSPLCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	502	23.5	931	2	T32919
2	484.5	22.6	738	2	hypothetical prote
3	398.5	18.6	741	2	dipeptidyl peptida
4	379	17.7	711	2	X-Pro dipeptidyl-p
5	359.5	16.8	793	2	dipeptidyl aminope
6	342.5	16.0	760	1	dipeptidyl-peptida
7	340.5	15.9	792	1	dipeptidyl-peptida
8	337.5	15.8	766	1	dipeptidyl-peptida
9	314	14.7	931	2	dipeptidyl aminope
10	310	14.5	818	1	dipeptidyl aminope
11	303.5	14.2	779	2	hypothetical prote
12	303.5	14.2	799	2	hypothetical prote
13	284.5	13.3	829	2	hypothetical prote
14	262	12.2	803	2	dipeptidyl aminope

15	248	11.6	803	2	I68600
16	248	11.6	865	2	I54331
17	245.5	11.5	795	2	F82858
18	244	11.4	743	2	T37700
19	226	10.6	631	2	H75007
20	209.5	9.8	683	2	E87495
21	207	9.7	642	2	C71137
22	188	8.8	709	2	B82580
23	185	8.6	759	2	I38593
24	171	8.0	591	2	H72474
25	168	7.9	622	2	F71174
26	167	7.8	632	2	E75057
27	152.5	7.1	657	2	E70025
28	151.5	7.1	569	2	S74053
29	149	7.0	536	2	F90299
30	147	6.9	721	2	T09631
31	146	6.8	608	2	F83397
32	143	6.7	828	2	G87584
33	140.5	6.6	572	2	F72455
34	140	6.5	732	1	JC4655
35	138	6.4	732	1	S07624
36	137.5	6.4	667	2	A87711
37	137	6.4	265	2	B84063
38	135.5	6.3	598	2	F84199
39	127	5.9	618	2	D87651
40	126.5	5.9	582	2	D72636
41	126.5	5.9	732	1	JU0132
42	123.5	5.8	411	2	AG1816
43	121.5	5.7	891	2	G82543
44	120	5.6	674	2	B84381
45	119	5.6	761	2	S44807

ALIGNMENTS

RESULT 1

T32919
hypothetical protein K02F2.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C:Accession: T32919
R:Maggi, L.; Goela, D.
submitted to the EMBL Data Library, January 1998
A:Description: The sequence of C. elegans cosmid K02F2.
A:Reference number: Z21246
A:Accession: T32919
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-931 <MAG>
A:Cross-references: UNIPROT:O44987; UNIPARC:UPI0000076BD8; EMBL:AF043699; FIDN:AAB97564.1
A:Experimental source: strain Bristol N2; clone K02F2
C:Genetics:
A:Gene: CESP:K02F2.1
A:Map position: 1
A:Introns: 58/3; 82/2; 131/2; 178/2; 275/3; 322/2; 404/2; 441/1; 464/2; 486/3; 528/2; 575/2

Alignment Scores:				
Pred. No.:	5.13e-39	Length:	931	
Score:	502.00	Matches:	122	
Percent Similarity:	48.4%	Conservative:	56	
Best local Similarity:	33.2%	Mismatches:	104	
Query Match:	23.5%	Indels:	86	
DB:	2	Gaps:	14	

US-10-825-632-4 (1-1197) x T32919 (1-931)

QY	3	TTTGAAGGCACCAAGACTCCCTTAGAGCATCCTTAGCTAGTCAGTACGTAAT	62
Db	563	TyrValAlaAsnGluSerHisProThrGluTrpAsnIle---CysValSerHisTyrArg	581
QY	63	CCTGGAGAGGTGACAGCGTACTGACCTGCTACTCATTCTTGTCATCAGTACAG	122
Db	582	ThrGlyGlnHisAlaGlnLeuThrGluSerGlyIle-----	593

```
QY 123 CACTGTGACTCTTTTATAGTAAGTATAGTATACCAAGAG-----AATCCACAC 170
Db 594 ---Cys-----PhetySerGluArgAlaAenGlyLysLeuAlaLeuAsePLeuAsePhis 610
QY 171 TGTGTGTGCTTACAGCTATCA-----ATCTCTGAAGATGACCCCACTTGGCAAAACA 224
Db 611 GlyPheAlaCysTyrMetThrSerValGlySerProAlaGlu-----CysArgPhe 627
QY 225 AAGGAATTT---TGG-----GCCACCAT 245
Db 628 TyrSerPheArgTrpLysGluAenGluValLeuProSerThrValTyrAlaAlaAseNile 647
QY 246 TTGGAATTCAGCAGGTCTCT---CTTCTCTGAC-----TATACTCTCTCCAGAAATTTTCTCT 296
Db 648 ThrValSerGlyHisProGlyGlnProAsePLeuHisPheAsePProGluMetIleGlu 667
QY 297 TTGAAAGT---ACTACTGTGATTTACATTTATGGATGCTCTACAGCCTCATGACTA 353
Db 668 PheGlnSerLysLysThrGlyLeuMetHisTyrAlaMetIleLeuAArgProSerAenPhe 687
QY 354 CAGCTCGGAAGAAATATCTACTGTGCTGTTCATATATGTGTGCTCT----- 401
Db 688 AspProTyrLysLysTyrProValPheHisTyrValTyrGlyGlyProGlyIleGlnIle 707
QY 401 ----- 401
Db 708 ValHisAsnAspPheSerTrpIleGlnTyrIleAArgPheCysArgLeuGlyTyrValVal 727
QY 401 ----- 401
Db 728 ValPheIleAspAenArgGlySerAlaHisArgGlyIleGluPheGluAArgHisIleHis 747
QY 402 -----CAGGTCAAATAGAATTCAGCATCAGGTGGAAGCACTCCAATATCTAGCTTCT 455
Db 748 LysLysMetGlyThrValGluValGluAsePLeuGlnValGlyLeuGlnMetLeuAlaGlu 767
QY 456 CGATAT---GATTTCATTGACTTAGATCGTGTGGGCATCCACGGCTGCTCTATGGAGGA 512
Db 768 ArgThrGlyGlyPheMetAsePTrpArgLeuTyrAsePValValHisGlyTrpSerTyrGlyGly 787
QY 513 TACTCTCCCTGTAGTCATTAATGACAGAGGTACAGATATCTTCAGGGTTGCTATTGCTGGG 572
Db 788 TyrMetAlaLeuGlnMetIleAlaLysHisProAenIleTyrArgAlaAlaIleAlaGly 807
QY 573 GCCCCAGTCACTCTGTGGATCTTATGATACAGGATACACGGAACGTTATATGGGTAC 632
Db 808 GlyAlaValSerAsePTrpArgLeuTyrAsePThrAlaTyrThrGluAArgTyrMetGlyTyr 827
QY 633 CCTGACCAGAAATCAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAGTTC 692
Db 828 Pro---LeuGluGluHisValTyrGlyAlaSerSerIleThrGlyLeuValGluLysLeu 846
QY 693 CCCTCTGACCAATCGTTTACTGCTCTTACATGTTTCTCGATGAGATGTCCTCATTTT 752
Db 847 ProAsePLeuProAsePLeuMetLeuValHisGlyLeuMetAsePLeuAenValHisPhe 866
QY 753 GCACATACAGTATATTACTAGTCTTTTATGAGGGCTGGAAGCCATATGATTTACAG 812
Db 867 AlaHisLeuThrHisLeuValAsePLeuGlyCysIleLysLysGlyLysTrpHisGluLeuVal 886
QY 813 ATCTATCTCAGAGAGACACAGCATAGAAGTTCCTGAATCGGGAGAACATATGAACTG 872
Db 887 IlePheProAenGluArgHisGlyValArgAenAsePAlaSerIleTyrLeuAsePAla 906
QY 873 CATCTTTTGACTTACCTTCAAGAA 896
Db 907 ArgMetMetTyrPheAlaGlnGln 914
```

RESULT 2
A87516
dipeptidyl peptidase IV [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus

C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C:Accession: A87516
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.I.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: A87516
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-738 <SFO>
A:Cross-references: UNIPROT:Q9A6E0; UNIPARC:UPI00000C7616; GB:AE005673; NID:g13423647; P
C:Genetics:
A:Gene: CC2154

Alignment Scores:
Pred. No.: 2,22e-37 Length: 738
Score: 484.50 Matches: 106
Percent Similarity: 48.3% Conservative: 53
Best Local Similarity: 32.2% Mismatches: 109
Query Match: 22.6% Indels: 61
DB: Gaps: 6

US-10-825-632-4 (1-1197) x A87516 (1-738)

```
QY 3 TTTGAGGACCAAAAGACTCCCTTTAGAGCATCACCTGTAGTGTAGTGTAGTGTAGTAAAT 62
Db 394 PheSerAlaSerIleAsePThrProIleGluArgArgLeuTyrGluValSerTyrAlaLys 413
QY 63 CTGTGAGAGGTGACAAAGGTGACTGACCGTGGCTACTCACATTTCTTGTGTGCATCAGTCAG 122
Db 414 ProGlyLysProLysAlaLeuThrSerAlaGlyGlyTyrTrpAlaAlaLysValAlaAseP 433
QY 123 CACTGTGACTCTTTTATAAGTATAGTATAACCAAGAAGATCCACACTGTGTGTCCCTT 182
Db 434 AsnGlyGlyAlaPheAlaGlyThrTyrSerAsePProLysThrProSerGlnThrAlaLeu 453
QY 183 TACAAGTACTCAAGTCTCT-----GAAGATGACCCCACTTGCAAAACAAAG 227
Db 454 TyrSerAlaAsePglyLysArgValArgTrpIleGluGluAseNlysLeuAlaGluGlyHis 473
QY 228 GAATTTTGGGCCACCATTTTGGATTGACGAGGTCTCTTCTCTGACTATATCTCTCCAGAA 287
Db 474 ProTyrTrp-----ProTyrAlaAlaAseNLeuProGln-----ProGlu 486
QY 288 ATTTTCTCTTTGAAAAGTACTACTGTGATTTACATTTATGGATGCTCTACAAGCCTCAT 347
Db 487 PheGlySerLeuLysAlaAlaAsePglyGluThrLeuHisTyrGluIleLeuLysProIle 506
QY 348 GATCTACAGCTGGAAGAATAATCTCTACTGTGTGTTTCATATATGGTGGTCTCAG--- 404
Db 507 GlyPheAsePProAlaLysLysTyrProAlaIleValSerValTyrGlyGlyProHisAla 526
QY 404 ----- 404
Db 527 GlnArgValMetLysAseNTrpHisSerProSerGluAArgThrTyrLeuGluAlaGlyTyr 546
QY 404 ----- 404
Db 547 ValIlePheLysLeuAsePAsnArgGlySerGlyAseNArgSerAlaLysPheMetArgAla 566
QY 405 -----GGTCAATAGAAATTCAGATCAGGTGGAAGGACTCCAATATCTA 449
Db 567 LeuAsePArgLysLeuGlyThrValGluValGluAsePLeuLeuLeuAlaLysPheLeu 586
QY 450 GCTTCTCGATATGATTTCAATTCACTTAGATCGTGTGGGCATCCACGGCTGCTCTATGGA 509
Db 587 AlaSerGln---ProTyrValAsePAlaAsePlysLeuGlyValMetGlyTrpSerTyrGly 605
QY 510 GGATACCTCTCCCTGATGGCATTAATGCAGAGGTGCAGATATCTCTCAGGGTTGCTATTGCT 569
Db 606 GlyPheMetAlaLeuMetLeuLeuThrAlaGluAseNThrProPheLysAlaGlyAlaAla 625
```

```
QY 570 GGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACCGAAGCTTTATATGGGT 629
|||||
Db 626 GlyAlaProThrGluTrpSerLeuThrAlaValProGluArgTyrMetGly 645
|||||
QY 630 CACCTGACACAGATGAACAGGCTATTACTTAGGATCTGTGGCCATGCACGACAGAAAG 689
|||||
Db 646 LysProAspGluAenLysAlaGlyTyrAlaTyrSerAspAlaAsnArgIleAspLys 665
|||||
QY 690 TTCCCTCTGAACCAAAATCGTTTACTCTCTACATGGTTTCTCGATGAGAAATGCCAT 749
|||||
Db 666 Leu-----AlaProGlySerLeuLeuLeuLeuHisGlyMetAlaAspAsnValIle 683
|||||
QY 750 TTGACATACACAGTATATTACTGAGTGTGTTTATAGTCAGGCTGGAAAGCCATATGTTA 809
|||||
Db 684 PheGluAenSerThrArgLeuMetAlaAlaLeuGlnArgLysAlaIleLeuPheGluMet 703
|||||
QY 810 CAGATCTCTCTCAGGAGACACAGC 836
|||||
Db 704 AlaMetTyrProGlyGluArgHisSer 712
|||||
RESULT 3
JC5142
X-Pro dipeptidyl-peptidase (EC 3.4.14.11) - Xanthomonas maltophilia
C:Species: Xanthomonas maltophilia
C>Date: 31-Jan-1997 #sequence_revision 27-Feb-1997 #text_change 09-Jul-2004
C:Accession: JC5142
R:Kabashima, T.; Ito, K.; Yoshimoto, T.
J. Biochem. 120, 1111-1117, 1996
A:Title: Dipeptidyl peptidase IV from Xanthomonas maltophilia: Sequencing and expression
A:Reference number: JC5142; MUID:97164011; PMID:9010758
A:Accession: JC5142
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-741 <KAB>
A:Cross-references: UNIPROT:P95782; UNIPARC:UPI0000085F8A; DDBJ:D83263; NID:g1753196; PI
C:Comment: This enzyme catalyzes the hydrolytic removal of amino-terminal dipeptidyl res
C:Superfamily: dipeptidyl-peptidase IV
F:4-18/Domain: transmembrane #status predicted <TM>
F:610/Active site: Ser #status predicted
F:685/Active site: Asp #status predicted
F:717/Active site: His #status predicted
Alignment Scores:
Pred. No.: 3,06e-29 Length: 741
Score: 398.50 Matches: 98
Percent Similarity: 44.3% Conservative: 57
Best Local Similarity: 28.0% Mismatches: 120
Query Match: 18.6% Indels: 75
DB: 2 Gaps: 8
US-10-825-632-4 (1-1197) x JC5142 (1-741)
QY 3 TTTGAGGACCAAGAAGACTCCCTTTAGACATCACTCTGATAGTTCAGTTACGTAAAT 62
|||||
Db 396 PheArgAlaGlyIleGluSerAlaArgGluSerGlnIleTyrAlaValPro---LeuGln 414
|||||
QY 63 CTGGAGAGGTGACAGGCTGACGTGACCGGTGCTACTCATCTCTGTCGATCAGTCAG 122
|||||
Db 415 GlyGlyGlnProGlnArgLeuSerLysAlaProGlyMetHisSerAlaSerPheAlaArg 434
|||||
QY 123 CACTGTGACTCTTTTATAAGTAAAGTATAGTAACCAAGAAATCCACACTGTGTGCCCTT 182
|||||
Db 435 AenAlaSerValTyrValAspSerTrpSerAsnAsnSerThrProGlnIleGluLeu 454
|||||
QY 183 TACAGCTATCAATGCTCTGAAGATGACCCAACTTGCAAAACAAAGGAATTTGGGCCACC 242
|||||
Db 455 PheArgAlaAenGly-----GluLysIleAlaThr 464
|||||
QY 243 ATTTTG-----GATTGAGGAGTCTCTCTTCCTGACTATACT- 278
|||||
Db 465 LeuValGluAenAspLeuAlaAspProLysHisProTyrAlaArgTyrArgGluAlaGln 484
|||||
```

```
QY 279 ---CCTCAGCAAAATTTTCTCTTTGAAAGTACTACTCGATTTTACA---TTGTATGGGATG 332
|||||
Db 485 ArgProValGluPheGlyThrLeuThrAlaAlaAspGlyLysThrProLeuAsnTyrSer 504
|||||
QY 333 CTCACAGCTCATGATCTACAGCTCGGAAAGAAATATCTCTGCTGTGTTTCATATAT 392
|||||
Db 505 ValIleLysProAlaGlyPheAspProAlaLysArgTyrProValAlaValTyrValTyr 524
|||||
QY 393 GGTGTCTCT----- 401
|||||
Db 525 GlyGlyProAlaSerGlnThrValThrAspSerTrpProGlyArgGlyAspHisLeuPhe 544
|||||
QY 401 ----- 401
Db 545 AsnGlnTyrLeuAlaGlnGlnGlyTyrValValPheSerLeuAspAsnArgGlyThrPro 564
|||||
QY 402 -----CAGGTCAAAATAGAAATGAC 422
|||||
Db 565 ArgArgGlyArgAspPheGlyAlaLeuTyrGlyLysGlnGlyThrValGluValAla 584
|||||
QY 423 GATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTCACTTACCTTAGCT 482
|||||
Db 585 AspGlnLeuArgGlyValAlaIleTrpLeu---LysGlnGlnProTrpValAspProAlaArg 603
|||||
QY 483 GTGGCATCCACGGCTGCTCTATGGAGGATACCTCTCCCTGATGGCATTAAATGCAGAGG 542
|||||
Db 604 IleGlyValGlnGlyTyrSerAsnGlyGlyTyrMetThrLeuMetLeuLeuAlaLysAla 623
|||||
QY 543 TCAGATATCTTCAGGGTGTCTATGCTGGGCCCCAGTCACTCTGTGGATCTTCTATGAT 602
|||||
Db 624 SerAspSerTyrAlaCysGlyValAlaGlyValAlaProValThrAspTrpGlyLeuTyrAsp 643
|||||
QY 603 ACAGGATACAGGAACGTTATATGGGTCACTCCCTGACAGAAATGCAACAGGGCTTACTTGA 662
|||||
Db 644 SerHisTyrThrGluArgTyrMetAspLeuProAlaArgAsnAspAlaGlyTyrArgGlu 663
|||||
QY 663 GGATCTGTGGCCATGCAAGCAGCAAGAAAGTTCCTCTGTAACCAAAATCGTTTACTGCTCTTA 722
|||||
Db 664 AlaArgValLeuThrHisIleGluGlyLeuArgSerPro-----LeuLeuLeuLeu 680
|||||
QY 723 CATGTTTCTCGATGATGATGTCATTTCACATACACCATATATTACTGATGTTTTTA 782
|||||
Db 681 HisGlyMetAlaAspAspAsnValLeuPheThrAsnSerThrSerLeuMetSerAlaLeu 700
|||||
QY 783 GTGAGGGCTGCAAGCCATATGATTTACAGATCTATCTCAGAGAGACACAGCATAGA 842
|||||
Db 701 GlnLysArgGlyGlnProPheGluLeuMetThrTyrProGlyAlaLysHisGlyLeuSer 720
|||||
QY 843 GTTCTCTGAATCGGAGAACATTTGAACTG 872
|||||
Db 721 GlyAlaAspAlaLeuHisArgTyrArgVal 730
|||||
RESULT 4
S66261
X-Pro dipeptidyl-peptidase (EC 3.4.14.11) - Flavobacterium meningosepticum
C:Species: Flavobacterium meningosepticum
C>Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C:Accession: S66261
R:Kabashima, T.; Yoshida, T.; Ito, K.; Yoshimoto, T.
Arch. Biochem. Biophys. 320, 123-128, 1995
A:Title: Cloning, sequencing, and expression of the dipeptidyl peptidase IV
A:Reference number: S66261; MUID:95314307; PMID:7793970
A:Accession: S66261
A:Status: preliminary
A:Molecule type: DNA
A:Cross-references: UNIPROT:Q47900; UNIPARC:UPI0000082115; EMBL:D42121; NID:g577283; PID:
C:Superfamily: dipeptidyl-peptidase IV
C:Keywords: dipeptidylpeptide hydrolase
Alignment Scores:
Pred. No.: 2.12e-27 Length: 711
```

Score: 379.00 Matches: 90
Percent Similarity: 43.8% Conservatives: 58
Best Local Similarity: 26.6% Mismatches: 118
Query Match: 17.7% Indels: 72
DB: 2 Gaps: 5

US-10-825-632-4 (1-1197) x S66261 (1-711)

```
QY 57 GTAATCCTCGAGAGTGCACAGGCTGACTGACCTGGCTACTCACATCTCTGTGTCATC 116
Db 390 IleaThrGlyLysThrGlnLeuLeuSerAenAlaGluGlyAenAenSerAlaAlaPhe 409
QY 117 AGTCAGCAGCTGCTCTCTTTATAAGTAAGTAGTAAACCAGAGAATCCACAGCTGGTG 176
Db 410 SerLysThrPheAenThrPheIleAenThrSerSerThrAlaAlaValProThrLysTyr 429
QY 177 TCCTTTTACAAGCTATCAAGTCTCTGAA-----GATGCCCAACT 215
Db 430 IleLeuLysAspAlaAenGlyLysAspValLysGluLeuGlnAenAenAspAspLeuLeu 449
QY 216 TCGAAACAAGAAATTTGGGCCACCATTTTGGATTTCAGCAGGTCTCTCTCTGACTAT 275
Db 450 AenLysLeuLysSer-----AspAenPhe 457
QY 276 ACTCCTCCAGAAATTTCTCTTTGAAAGTACTACTGGATTTCATTTGATGGAGTCTC 335
Db 458 IleAlaLysGluPheIleThrIleProAenAlaAlaGlyAspGlnMetAenAlaTrpMet 477
QY 336 TACAAGCCTCATGCTTACAGCCTGGAAAGAAATATCTCTACTGTGCTGTTTCATATPGGT 395
Db 478 IleLysProLysAenPheAspProAlaLysLysTyrProValPheMetPheGlnTyrSer 497
QY 396 GGTCTCT----- 401
Db 498 GlyProGlySerGlnGlnValAlaAenSerTrpAspGlyGlyAenGlyIleTrpPheAsp 517
QY 401 ----- 401
Db 518 MetLeuAlaGlnLysGlyTyrLeuValValCysValAspGlyAenGlyThrGlyPheArg 537
QY 402 -----CAGGGTCAAAATAGAAATGACGATCAG 428
Db 538 GlyThrLysTyrLysLysValThrTyrLysAenLeuLysTyrGluIleGluLysGln 557
QY 429 GTGGAAGGACTCCAATATCTAGCTCTCGATATGATTTTCATTTGATTCGATCGGTGGC 488
Db 558 IleThrAlaAlaLysTrpLeuGlyAenGln---SerTyrValAspLysSerArgIleGly 576
QY 489 ATCCACGGCTGCTCTATGAGGATACCTCTCCCTGATGGCATTAATGACAGAGTCAGAT 548
Db 577 IlePheGlyTrpSerTyrGlyTyrMetAlaSerLeuAlaMetThrLysGlyAlaAsp 596
QY 549 ATCTTCAGGTTGCTATTGCTGGGGCCAGTCACCTCTGTGATCTTCTATGATACAGGA 608
Db 597 ValPheLysMetGlyIleAlaValAlaProValThrAenThrPheGlyAspSerIle 616
QY 609 TACACGAAAGCTTATATGGGTCACTCTGACCAAGATGAACAGGGCTATTACTTAGGATCT 668
Db 617 TyrThrGluArgPheLeuGlnThrProGlnGluAenLysAspGlyTyrAspLeuAenSer 636
QY 669 GTGCCCATGCAAGCAAGAAAGTTCCCTCTCTGAAACCAATCGTTTACTCTCTTACATGGT 728
Db 637 ProThrThrTyrAlaLysLeuLeu-----LysGlyLysPheLeuLeuIleHisGly 653
QY 729 TTCCTGATGAGATGTCCATTTTGCACATACAGTATATTTACTGAGTTTTTTAGTGAGG 788
Db 654 ThrAlaAspAenValHisPheGlnAenSerMetGluPheSerGluAlaLeuIleGln 673
QY 789 GCTGGAAGCCATATGATTTACAGATCTATCTCAGGAGAGACACAGCATTAAGAGTTTCCT 848
Db 674 AenLysLysGlnPheAspPheMetAlaTyrProAspLysAenHisSerIleIleGlyGly 693
QY 849 GAATCGGGAGAACATTATGAAGTGCATCTTTTGCACCTACCTTCAAGAAACCTT 902
```

Db 694 AenThrArgProGlnLeuTyrGluLysMetThrAenTyrIleLeuGluAenLeu 711
RESULT 5
T41703
dipeptidyl aminopeptidase - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: T41703
R:Murphy, L.; Harris, D.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
Submitted to the EMBL Data Library, August 1998
A:Reference number: Z22011
A:Accession: T41703
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-793 <MUR>
A:Cross-references: UNIPROT:O14073; UNIPARC:UPI000013AB95; EMBL:AL031180; PIDN:CAA20138.1
A:Experimental source: strain 972h-; cosmid c2E11 -chimeric
C:Genetics:
A:Gene: SPAC2E11.08
A:Map position: 1
C:Superfamily: dipeptidyl-peptidase IV

Alignment Scores:

Pred. No.:	1-52e-25	Length:	793
Score:	359.50	Matches:	113
Percent Similarity:	43.5%	Conservative:	43
Best Local Similarity:	31.5%	Mismatches:	128
Query Match:	16.8%	Indels:	75
DB:	2	Gaps:	17

US-10-825-632-4 (1-1197) x T41703 (1-793)

```
QY 15 AAAGACTCCCTTTAGACATCACCTGTAGTACGTAGTAAATCTCGGAGAGTG 74
Db 442 LysAspSer---ThrGluArgHisLeuTyrTyrValSer---LeuAspThrLeuGluIle 459
QY 75 ACAAGGCTGACTGACCGTGGCTACTCA-----CATTCCTGCTGATCAGTCAGCAC 125
Db 460 TyrGlyIleThrAspAenGlyGluAspGluGlyTyrTyrSerThrSerPheSerPhe 479
QY 126 TGTCACTCTTTTATAGTAAGTAACTAGTAAACACAGAAAGATCCACACTGTGTCTCCCTTTAC 185
Db 480 GlyAspPheTyrValLeuAenTyrHisGlyProAspValPro-----TrpGln 495
QY 186 AAGCTATCAAGTCTCTGAAGATGACCCCACTGCAAAACAAAGAAATTTTGGCCACCAT 245
Db 496 GluLeuArgSerThrLysAspLysAspTyrCysLeuSerLeuGluThrAenSerArgLeu 515
QY 246 TTGGATTTCAGCAGTCTCTCTCTGACTATCTCTCCAGAAAT---TTT 293
Db 516 LysGlnGln-----LeuSerSerIleThrLeuProSerValGluTyrGlyLysLeu 532
QY 294 TCTTTTGAAGTACTACTGGATTTTACATTTGATGGGATGCTCTACAAGCCTCATGATCTA 353
Db 533 ThrPheAenAspThrThr---PheAenPhe-----MetGluArgArgProAenAspPhe 549
QY 354 CAGCCTGGAAGAAATATCTCTACTGTGCTTTCATATATATGTTGCTCAGGGTCAA--- 410
Db 550 AspValAenLysLysTyrProValLeuPheAlaTyrGlyProGlySerGlnGln 569
QY 410 ----- 410
Db 570 ValAlaLysLeuPheArgValAspPheGlnAlaTyrLeuAlaSerHisProAspPheGlu 589
QY 411 -----ATAGAAATTTGACGATCAGGTGGGAAGGA-----CTCCAATAT 446
Db 590 PheIleValValThrLeuAspGlyArgGlyThrGlyPheAenGlyAenAlaPheArgTyr 609
QY 447 CTAGCTTCTCGA-----TATGAT----- 464
Db 610 SerValSerHisLeuGlyGluTrpGluSerTyrAspGlnGlyGlnAlaGlyLysPhe 629
```

```
QY 465 -----TTTCATTGACTTAGATCGTGGGCATCCACGGCTGGTCTCTATGGA 509
Db 630 TrpAlaAspLeuProPheValAspGluAsnHisValGlyIleTrpGlyTrpSerTyrgly 649
QY 510 GGATACCTCTCCCTGAGGATTAAGTACAGAGGTGAGATATCTTCAGGGTTCCTATGCT 569
Db 650 GlyTyrLeuThrLeuLysThrLeu---GluThrGlnAspValPheSerTyrglyMetAla 668
QY 570 GGGGCCCCAGTCACTCTGTCGATCTTCTATGATACAGGATACACGGAACGTTATGGGT 629
Db 669 ValAlaProValThrAspTrpArgLeuTyrAspSerValTyrThrGluArgTyrMetAsp 688
QY 630 CACCTGACAGGATGAACAGGCTATTACTTAGGATCTGTGGCCATCAAGCAGAGAAAG 689
Db 689 LeuProGlnTyrAsnLysGluGly---TyrLysAsnSerGlnIleHisAspTyrGluLys 707
QY 690 TTCCCTCTGAACCAATCCCTTACTGCTCTTACATGGTTTCTCGATGAGAATGTCCAT 749
Db 708 Phe---LysGlnLeuLysArgPhePheValAlaHisGlyThrGlyAspAsnValHis 726
QY 750 TTTCACATACCATATPATTACTGAGTGTCTTTTGTAGTGGGCTGGA---AAGCCATATCAT 806
Db 727 PheGlnHisSerMetHisLeuMetAspGlyLeuAsnLeuAlaAsnCysTyrAsnTyrAsp 746
QY 807 TTACAGATCTATCTCAGGAGACACAGCATAGAAGTTCCTGAATCGGAGAACATAT 866
Db 747 MetAlaValPheProAspSerAlaHisSerIle-----SerTyrHisAsnAla 762
QY 867 GAATCGATCTTTTGGACTACCTTCAAGAAACCTTGGATCAGTATTGCTGCTCTA 923
Db 763 SerLeuSerIleTyrHisArgLeuSerGluTrpIleGlyAspAlaLeuGlyArgIle 781

RESULT 6
S23752
dipeptidyl-peptidase IV (EC 3.4.14.5) alpha chain - mouse
N:Alternate names: CD26 alpha subunit; THAM alpha subunit
C:Species: Mus musculus (house mouse)
C>Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: S23752; A46465; A56030
R:Marguet, D.; Bernard, A.M.; Vivier, I.; Darmoul, D.; Naquet, P.; Pierres, M.
J. Biol. Chem. 267, 2200-2208, 1992
A:Title: cDNA cloning for mouse thymocyte-activating molecule. A multifunctional ecto-di
A:Reference number: S23752; MUID:92129288; PMID:1370813
A:Accession: S23752
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-760 <MAR>
A:Cross-references: UNIPARC:UPI0000172A2D; EMBL:X58384
R:Vivier, I.; Marguet, D.; Naquet, P.; Bonicel, J.; Black, D.; Li, C.X.; Bernard, A.M.;
J. Immunol. 147, 447-454, 1991
A:Title: Evidence that thymocyte-activating molecule is mouse CD26 (dipeptidyl peptidase
A:Reference number: A46465; MUID:91302787; PMID:1712807
A:Accession: A46465
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-20 <VIV>
A:Cross-references: UNIPARC:UPI0000172A2E
A:Experimental source: M14.T thymoma cells, Swiss nu/nu
A>Note: sequence extracted from NCBI backbone (NCBI:42236)
R:Bernard, A.M.; Mattei, M.G.; Pierres, M.; Marguet, D.
Biochemistry 33, 15204-15214, 1994
A:Title: Structure of the mouse dipeptidyl peptidase IV (CD26) gene.
A:Reference number: A56030; MUID:95032780; PMID:7999781
A:Accession: A56030
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 746-760 <BER>
A:Cross-references: UNIPARC:UPI0000172A2F; GB:U12620
C:Genetics:
C:Superfamily: dipeptidyl-peptidase IV
C:Keywords: dipeptidylpeptide hydrolase; glycoprotein; transmembrane protein
F:213,223,315,514,679/Binding site: carbohydrate (Asn) (covalent) #status predicted
```

```
F:624,702,734/Active site: Ser, Asp, His #status predicted

Alignment Scores: 6.13e-24 Length: 760
Pred. No.: 342.50 Matches: 101
Score: 42.1% Conservative: 51
Best Local Similarity: 28.0% Mismatches: 120
Query Match: 16.0% Indels: 89
DB: 1 Gaps: 14

US-10-825-632-4 (1-1197) x S23752 (1-760)

QY 15 AAAGACTCCCTTTAGAGCATCACCTGTAGTAGTCAGTTACGTAAATCCTGGAGAGGTG 74
Db 417 LysGluMetProGlyGlyArgAsnLeuTyrLysIle----- 428
QY 75 ACAAGGTGACTGACCGTGGCTACTCACATTTCTTGTGCATCAGT----- 119
Db 429 ---GlnLeuThrAsp-----HisThrAsnValLysCysLeuSerCysAspLeuAsnPro 445
QY 120 CAGCAGCTGCACTTCTTTATAGTAAGTATAGTAACCAAGAAATCCACAC----- 170
Db 446 GluArgCysGlnTyrTyrAlaValSerPheSerLysGluAlaLysTyrTyrGlnLeuGly 465
QY 171 TGT-----GTGTCCCTTTACAAGCTATCATCAAGTCTCGAAGATGACCCAACTTGC 218
Db 466 CysTrpGlyProGlyLeuProLeuTyrThrLeuHisArgSerThrAspHisLysGluLeu 485
QY 219 AAACAAGGAAATTTGGGCCACCATTTGGATTGATCAGCAGGT-----CCTCTTCT 269
Db 486 ArgValLeuGlu-----AspAsnSerAlaLeuAspArgMetLeuGln 499
QY 270 GACTATACTCTCCAGAA-----ATTTCTCTTTTGAAGTACTACTGGATTACA 320
Db 500 AspValGlnMetProSerLysLysLeuAspPheIleValLeuAsnGluThrArgPhe--- 518
QY 321 TTGTATGGGATGCTTACAAAGCCTCATGATCTACAGCCTGGAAGAAATATCTACTGTG 380
Db 519 TrpTyrGlnMetIleLeuProProHis---PheAspLysSerLysLysTyrProLeuLeu 537
QY 381 CTGTTCATATATGGTGGTCTCT----- 401
Db 538 LeuAspValTyrAlaGlyProCysSerGlnLysAlaAspAlaSerPheArgLeuAsnTrp 557
QY 401 ----- 401
Db 558 AlaThrTyrLeuAlaSerThrGluAsnIleIleValAlaSerPheAspGlyArgGlySer 577
QY 402 -----CAGGGTCAATAGAAATT 419
Db 578 GlyTyrGlnGlyAspLysIleMetHisAlaIleAsnArgArgLeuGlyThrLeuGluVal 597
QY 420 GACGATCAGGTGGAAGGACTCCCAATATCTAGCTTCTCGATATGATTTTCATTGACTTAG 479
Db 598 GluAspGlnIleGluAlaAlaArgGlnPheVal---LysMetGlyPheValAspSerLys 616
QY 480 CGTGTGGGCATCCACGGCTGCTTATGAGGATACCTCTCCCTGATGCGATTATGCGAG 539
Db 617 ArgValAlaIleTrpGlyTrpSerTyrGlyGlyTyrValThrSerMetValleuGlySer 636
QY 540 AGGTGAGATATCTTCAGGGTGTCTATGTGTGGGGCCCCAGTCCTCTGTGATCTTCTAT 599
Db 637 GlySerGlyValPheLysCysGlyIleAlaValAlaProValSerArgTrpGluTyrTyr 656
QY 600 GATACAGGATACACGGAACGTTATATGGGT-----CACCTGTGACAGAAATGAACAGGGC 653
Db 657 AspSerValTyrThrGluArgTyrMetGlyLeuProIleProGluAspAsnLeuAspHis 676
QY 654 TATTACTTAGGATCTGTGGCCATGCGAAGCAGAAAGTTCCTCCCTCGAACCAATCGTTTA 713
Db 677 TyrArgAsnSerThrValMetSerArgAlaGluHisPhe-----LysGlnValGluTyr 694
QY 714 CTGCTCTTACATGGTTCCTGGATGAGAATGTCCATTTTGGCACATACCATATATTACTG 773
```

Db 695 LeuLeuIleHisGlyThrAlaAspAspAenValHisPheGlnGlnSerAlaGlnIleSer 714
Qy 774 AGTTTTAGTGGGCTGGAAGCCATGATTTTACAGATCTATCTCTCAGGAGACAC 833
Db 715 LysValLeuValAspAlaGlyValAspPheGlnAlaMetTrpTyrThrAspGluAspHis 734
Qy 834 AGCATAGAGTCTCTGAATCGGAGAACATTATGAACTGCATCTTTTGCACACTCTCAA 893
Db 735 GlyIleAlaSerSerThrAlaHisGlnHisIleTyrSerHisMetSerHisPheLeuGln 754
Qy 894 GAA 896
Db 755 Gln 755

RESULT 7
A39914
dipeptidyl-peptidase IV (EC 3.4.14.5), membrane-bound form precursor - rat
N:Alternate names: GP110; membrane glycoprotein 110K; OX-61
N:Contains: dipeptidyl-peptidase IV, soluble form
C:Species: Rattus norvegicus (Norway rat)
C:Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text change 09-Jul-2004
C:Accession: A39914; A33315; B33315; A60730; A42203; S38949; A31781
R:Hong, W.; Doyle, D.
Proc. Natl. Acad. Sci. U.S.A. 84, 7962-7966, 1987
A:Title: cDNA cloning for a bile canalicular domain-specific membrane glycoprotein of rat
A:Reference number: A39914; MUID:88068516; PMID:3479775
A:Accession: A39914
A:Molecule type: mRNA
A:Residues: 1-792 <HON>
A:Cross-references: UNIPROT:P14740; UNIPARC:UPI000017098A; GB:J02997; NID:G204463; PIDN:
R:Ogata, S.; Misumi, Y.; Ikehara, Y.
J. Biol. Chem. 264, 3596-3601, 1989
A:Title: Primary structure of rat liver dipeptidyl peptidase IV deduced from its cDNA an
A:Reference number: A33315; MUID:89123496; PMID:2563382
A:Accession: A33315
A:Molecule type: mRNA
A:Residues: 1-37, 'A', 39-182, 'I', 184-331, 'T', 333-351, 'C', 353-393, 'V', 395-561, 'L', 563-623,
A:Cross-references: UNIPARC:UPI0000129841; GB:J04591; NID:G203973; PIDN:AAA41096.1; PID:
A:Note: the authors translated the codon GCG for residue 38 as Arg, ACC for residue 332
A:Accession: B33315
A:Molecule type: protein
A:Residues: 1-20, 35-54, 427-443, 505-509, 511-520, 530-538, 593-600, 602-608, 618-627 <HO2>
A:Cross-references: UNIPARC:UPI0000172A37
A36; UNIPARC:UPI0000172A37
R:McCaughan, G.W.; Wickson, J.E.; Creswick, P.F.; Gorrell, M.D.
Hepatology 11, 534-544, 1990
A:Title: Identification of the bile canalicular cell surface molecule GP110 as the ectop
quence.
A:Reference number: A60730; MUID:90228896; PMID:1970322
A:Accession: A60730
A:Molecule type: protein
A:Residues: 28-47, 'XX', 50-53, 55-58 <MCC>
A:Cross-references: UNIPARC:UPI0000172A38
R:Ogata, S.; Misumi, Y.; Teuji, E.; Takami, N.; Oda, K.; Ikehara, Y.
Biochemistry 31, 2582-2587, 1992
A:Title: Identification of the active site residues in dipeptidyl peptidase IV by affini
A:Reference number: A42203; MUID:92190188; PMID:1347701
A:Accession: A42203
A:Molecule type: protein
A:Residues: 'R', 625-630, 'X', 632-648 <OG2>
A:Cross-references: UNIPARC:UPI0000172A39
R:Iwaki-Egawa, S.; Watanabe, Y.; Fujimoto, Y.
Biol. Chem. Hoppe-Seyler 374, 973-975, 1993
A:Title: N-terminal amino acid sequence of the 60-kDa protein of rat kidney dipeptidyl p
A:Reference number: S38949; MUID:94128239; PMID:7905271
A:Accession: S38949
A:Status: preliminary
A:Molecule type: protein
A:Residues: 281-302 <IWA>
A:Cross-references: UNIPARC:UPI0000172A3A
R:Hong, W.; Doyle, D.
J. Biol. Chem. 263, 16892-16898, 1988

A:Title: Membrane orientation of rat gp110 as studied by in vitro translation.
A:Reference number: A31781; MUID:89034185; PMID:3182821
A:Accession: A31781
A:Molecule type: mRNA
A:Residues: 1-40 <HO3>
A:Cross-references: UNIPARC:UPI0000172A3B
C:Superfamily: dipeptidyl-peptidase IV
C:Keywords: dipeptidyl-peptidase, membrane-bound form #status experimental <MATM>
F:1-792/Product: dipeptidyl-peptidase, signal sequence #link MATS #status experimental <SIG>
F:1-28/Domain: intracellular #status predicted <INT>
F:1-6/Domain: transmembrane #status predicted <TMN>
F:7-28/Domain: extracellular #status predicted <EXT>
F:29-792/Domain: propeptide #link MATS #status experimental <PRO>
F:35-792/Product: dipeptidyl-peptidase, soluble form #status experimental <MATS>
F:83 90 148 217 227 319 521 686/Binding site: carboxylate (Asn) (covalent) #status pred
F:631/Active site: Ser #status experimental
F:709, 741/Active site: Asp, His #status predicted
Alignment Scores:
Pred. No.: 9,56e-24 Length: 792
Score: 340.50 Matches: 103
Percent Similarity: 42.9% Conservative: 51
Best Local Similarity: 28.7% Mismatches: 120
Query Match: 15.9% Indels: 85
DB: 1 Gaps: 15
US-10-825-632-4 (1-1197) x A39914 (1-792)
Qy 15 AAGACTCCCTTTAGAGCATCACCCTGTAGTACGTACGTAAATCTCGGAGAGTG 74
Db 424 LysGluMetProGlyGlyArgAsnLeuTyrLysIle 435
Qy 75 ACAAGGTGACTGACCGTGGCTACTCACATCTTCTGTCATCAGT 119
Db 436 ---GlnLeuThrAsp-----HisThrAsnLysLysCysLeuSerCysAspLeuAsnPro 452
Qy 120 CAGCAGCTGTGACTCTCTT-----ATAAGTAAGTAGTAGTAACAGAGAACATCCACAC 170
Db 453 GluArgCysGlnTyrTyrSerValSerLeuSerLysGluAlaLysTyrTyrGlnLeuGly 472
Qy 171 TGT-----GTGTCCTTTACAGCTATCAAGTCTCGAGAGTCTCTCTCGATATACT 218
Db 473 CysArgGlyProGlyLeuProLeuTyrThrLeuHisArgSerThrAspGlnLysGluLeu 492
Qy 219 AAAACAAGAAATTTGGGCCACCATTTTGGATTTCAGCAGGTCTCTCTCGATATACT 278
Db 493 ArgValLeuGluAsp---AsnSerAlaLeuAspLysMet-----LeuGlnAspValGln 509
Qy 279 CTTCCAGAA-----ATTTCTCTTTGAAAGTACTACTGGATTTACATTTGATGGG 329
Db 510 MetProSerLysLysLeuAspPheIleValLeuAsnGluThrArgPhe---TrpTyrGln 528
Qy 330 ATGCTCTACAGCCTCATGATCTACAGCCTGGAAAGAAATATCTCTACTGCTGCTCAT 389
Db 529 MetIleLeuProHis---PheAspLysSerLysLysTyrProLeuLeuIleAspVal 547
Qy 390 TATGGTGGTCTCT----- 401
Db 548 TyrAlaGlyProCysSerGlnLysAlaAspAlaAlaPheArgPheAsnTrpAlaThrTyr 567
Qy 401 ----- 401
Db 568 LeuAlaSerThrGluAsnIleIleValAlaSerPheAspGlyArgGlySerGlyTyrGln 587
Qy 402 -----CAGGCTCAAAATAGAAATGACGATCAG 428
Db 588 GlyAspLysIleMethHisAlaIleAsnLysArgLeuLeuGluValGluAspGln 607
Qy 429 GTGGAAGGACTC---CAATATCTAGCTTCTCGATATGATTTTCATTGACTTAGATCGGTG 485
Db 608 IleGluAlaAlaArgGlnPheLeu-----LysMetGlyPheValAspSerLysGlnVal 625

Db 551 CysSerGlnLysAlaAspThrValPheArgLeuAsnTrpAlaThrTyrLeuAlaSerThr 570
QY 401 -----
Db 571 GluAsnIleValAlaSerPheAspGlyArgGlySerGlyTyrGlnGlyAspLysIle 590
QY 402 -----CAGGTCAAATAGAAATGACGATCAGGTGGAGGACTC 440
Db 591 MetHisAlaIleAsnArgArgLeuGlyThrPheGluValGluAspGlnIleGluAlaAla 610
QY 441 CAATATCTAGCTCTTCGATATGATTTCAITGACTTAGATCGTGGGCGATCCACGGCTGG 500
Db 611 ArgGlnPhe---SerLysMetGlyPheValAspAsnLysArgIleAlaIleTrpGlyTrp 629
QY 501 TCCTATGGAGGATACCTCCCTGATGGCATTAAATGACAGAGTGCAGATATCTTCAGGGTT 560
Db 630 SerTyrGlyGlyTyrValThrSerMetValLeuLeuGlySerGlySerGlyValPheLysCys 649
QY 561 GCTATTCTGGGCGCCAGCTCTGCTGATCTTCTATGATACAGGATACACGGAACGT 620
Db 650 GlyIleAlaValAlaProValSerArgTrpGluTyrTyrAspSerValTyrThrGluArg 669
QY 621 TATATGGT-----CACCTGACACAGATGAACAGGCTATTACTAGGATCTGTGGCC 674
Db 670 TyrMetGlyLeuProThrProGluAspAsnLeuAspHisTyrArgAsnSerThrValMet 689
QY 675 ATGCAGCAGAAAGTTCCCTCTGACCAACATCGTTTACTGCTTACATGCTTCTTCCTG 734
Db 690 SerArgAlaGluAsnPhe-----LysGlnValGluTyrLeuLeuIleHisGlyThrAla 707
QY 735 GATGAGNATCTCCATTTGCATACAGTATATTTACTAGTATTTTACTGAGGCTGGA 794
Db 708 AspAspAsnValHisPheGlnSerAlaGlnIleSerLysAlaLeuValAspValGly 727
QY 795 AAGCATATGATTACAGATCTATCTCTCAGGAGACACACAGATAAGAGTCTCTGAATCG 854
Db 728 ValAspPheGlnAlaMetTrpTyrThrAspGluAspHisGlyIleAlaSerSerThrAla 747
QY 855 GGAGAACATATTGAACCTGATCTTTTGACACTACCTTCAAGAA 896
Db 748 HieGlnHisIleTyrThrHisMetSerHisPheIleLysGln 761
RESULT 9
A49737
dipeptidyl aminopeptidase (EC 3.4.14.-) - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein O5045; protein YOR219c; protein YOR50-9
C:Species: Saccharomyces cerevisiae
C:Date: 03-Mar-1994 #sequence revision 03-Mar-1994 #text_change 09-Jul-2004
C:Accession: A49737; S45451; S60946; S67112; S71721
R:Santa Anna-A, S.; Herskowitz, I.
submitted to the Protein Sequence Database, July 1993
A:Reference number: A49737
A:Accession: A49737
A:Molecule type: DNA
A:Residues: 1-931 <SAN>
A:Cross-references: UNIPROT:P33894; UNIPARC:UPI0000136060; GB:L21944; NID:g347196; PIDN:
R:Anna-Arriola, S.S.; Herskowitz, I.
Yeast 10, 801-810, 1994
A:Title: Isolation and DNA sequence of the STE13 gene encoding dipeptidyl aminopeptidase
A:Reference number: S45451; MUID:95066382; PMID:7975897
A:Accession: S45451
A:Molecule type: DNA
A:Residues: 1-931 <ANN>
A:Cross-references: UNIPARC:UPI0000136060; EMBL:L21944; NID:g347196; PIDN:AAA35119.1; PI
R:Gallison, F.; Dujon, B.
submitted to the EMBL Data Library, October 1995
A:Description: Sequence and analysis of a 33 kb fragment from the right arm of chromosome
A:Reference number: S60946
A:Accession: S60946
A:Molecule type: DNA
A:Residues: 1-931 <GAL>
A:Cross-references: UNIPARC:UPI0000136060; EMBL:X92441; NID:g1050762; PIDN:CAA63182.1; F
R:Boyer, J.; Fairhead, C.; Gallion, L.; Gallison, F.; Michaux, G.; Thierry, A.; Dujon, B.

submitted to the Protein Sequence Database, July 1996
A:Reference number: S67104
A:Accession: S67112
A:Molecule type: DNA
A:Residues: 1-931 <BOY>
A:Cross-references: UNIPARC:UPI0000136060; EMBL:275127; NID:g1420507; PIDN:CAA99437.1; PI
A:Experimental source: strain S288C
R:Gallison, F.; Dujon, B.
Yeast 12, 877-885, 1996
A:Title: Sequence and analysis of a 33 kb fragment from the right arm of chromosome XV of
A:Reference number: S71713; MUID:96437977; PMID:8840505
A:Accession: S71721
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-931 <GAW>
A:Cross-references: UNIPARC:UPI0000136060; EMBL:X92441; NID:g1050762; PIDN:CAA63182.1; PI
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1995
C:Genetics:
A:Gene: SGD:STE13; YC11
A:Cross-references: SGD:S0005745; MIPS:YOR219c
A:Map position: 15R
C:Function:
A:Description: involved in processing of alpha-factor prepropheromone
C:Keywords: dipeptidylpeptide hydrolase; glycoprotein; transmembrane protein
F:113-150/Domain: transmembrane #status predicted <TM>
F:377/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:785,863,896/Active site: Ser, Asp, His #status predicted
Alignment Scores:
Pred. No.: 3 18e-21 Length: 931
Score: 314.00 Matches: 100
Percent Similarity: 39.4% Conservative: 54
Best Local Similarity: 25.6% Mismatches: 133
Query Match: 14.7% Indels: 104
DB: 2 Gaps: 12
US-10-825-632-4 (1-1197) x A49737 (1-931)
QY 3 TTGGAAGGCCAACAGACTCCCTTTAGAGCATACCTGTACGTAGTCAGTTACGTA--- 59
Db 562 PheThrAlaAsnGluIleGlyValMetSerGlnHisLeuTyrSerIleSerLeuThrAsp 581
QY 60 -----AATCTGGAGAGGTGCACAAAGCTGACT 86
Db 582 SerThrThrGlnAsnThrPheGlnSerLeuGlnAsnPro-----Ser 595
QY 87 GACCGTGGCTACATCATTTCTGCTGCATCAGTACAGCAGCTGTGACTTCTTTTATAGTAAG 146
Db 596 AsplysTyrAspPheTyrAspPheGluLeuSerSerSerAlaArgTyrAlaIleSerLys 615
QY 147 TATAGTAACACAGAAATCCA----- 167
Db 616 LysLeuGlyProAspThrProIleLysValAlaGlyProLeuThrArgValLeuAsnVal 635
QY 168 -----CACGTGTGTCTCCCTTTACAGCTATCAAGTCTCAAGATGACCCAACTTGC 218
Db 636 AlaGluIleHisAspAspSerIleLeuGlnLeuThrLysAspGlu----- 650
QY 219 AAAACAAGAAATTTTGGGCCACCATTTTGGATTCCAGCAGGTCTCTCTCTGACTATACT 278
Db 651 LysPheLysGlu-----LysIleLysAsnTyrAsp 660
QY 279 CCTCCAGAAATTTCTTTTGAAAGTACTACT-----GGATTTTACATTGTATGGG 329
Db 661 LeuPro----IleThrSerTyrLysThrMetValLeuAspAspGlyValGluIleAsnTyr 679
QY 330 ATGCTCTACAGGCTCATCATCTACAGCTCGAAGAAATATCTCTGCTGCTGTTCTATA 389
Db 680 IleGluIleLysProAlaAsnLeuAsnProLysLysLysTyrProIleLeuValAsnIle 699
QY 390 TATGCTGCTCTCAGGGTCAA----- 410
Db 700 TyrGlyGlyProGlySerGlnThrPheThrLysSerSerLeuAlaPheGluGlnAla 719

Qy 411 -----ATAGAAATTGACGATCAGGTGGAGGA----- 437
Db 720 ValValSerGlyLeuAspValIleValLeuGlnIleGluProArgGlyThrGlyGlyLys 739
Qy 437 ----- 437
Db 740 GlyTrpSerPheArgSerTrpAlaArgGluLysLeuGlyTyrrTrpGluProArgAspIle 759
Qy 438 -----CTCCAATATCTAGCTTCTCGATATGATTTTCATTGACTTAGATCGGTG 485
Db 760 ThrGluValThrIleLysPheIleGlnArgAsnSerGlnHisIleAspGluSerLysIle 779
Qy 486 GGCATCAGCGCTGCTTATGAGATACCTCTCCCTGATGGCATTA---ATGCAGAGG 542
Db 780 AlaIleTrpGlyPheThrSerLeuLysThrValGluLeuLeuAspAsn 799
Qy 543 TCAGATATCTCAGGTTCTATTCTGGGGCCCGAGTCACTCTGTGATCTTCTATGAT 602
Db 800 GlyAspThrPheLysTyrrAlaMetAlaValAlaProValThrAsnTrpThrLeuTyrrAsp 819
Qy 603 ACAGGATACACGGAAGTATATGGTCACTGACAGATGAACAGGGCTATTACTTA 662
Db 820 SerValTyrrThrGluArgTyrrMetAsnGlnProSerGluAsnHisGluGlyTyrrPheGlu 839
Qy 663 GGAATCTGTGGCCATGCAAGCAGAAAGTTCCCTCTGAACCAAAATCGTTTACTGCTCTTA 722
Db 840 ValSerThrIleGlnAsnPheLysSerPheGluSer---LeuLysArgLeuPheIleVal 858
Qy 723 CATGGTTCTCTGGATGAGATGTCATTTTGCACATACCACTATATTACTGAGTTTCTTA 782
Db 859 HisGlyThrPheAspAspAsnValHisIleGlnAsnThrPheArgLeuValAspGlnLeu 878
Qy 783 GTGAGCGCTGGA---AAGCCATATGATTACAGATCTATCTCTCAGGAGACACAGCAT 839
Db 879 AsnLeuLeuGlyLeuThrAsnTyrrAspMetHisIlePheProAspSerAspHisSerIle 898
Qy 840 AGAGTCTCTGAAATCGGGAGAACATTATGACCTGATCTTTTGCATCTACCTCAAGAAAC 899
Db 899 ArgTyrrHisAsnAlaGlnArgIleValPheGlnLysLeuTyrrTyrrLeuArgAspAla 918
Qy 900 CTGGATCAGCATTTCTGCTCTAAAGTGATA 932
Db 919 PheAlaGluArgPheAspAsnThrGluValLeu 929
RESULT 10
A30107
dipeptidyl aminopeptidase B (EC 3.4.14.-) - yeast (Saccharomyces cerevisiae)
N/Alternate names: protein YHR028c
C/Species: Saccharomyces cerevisiae
C/Date: 07-Jun-1990 #sequence_revision 30-May-1997 #text_change 09-Jul-2004
C/Accession: S46780; A30107
R/Du, Z.
submitted to the EMBL Data Library, June 1994
A/Description: The sequence of *S. cerevisiae* cosmid 8082.
A/Reference number: S46773
A/Accession: S46780
A/Molecule type: DNA
A/Residues: 1-818 'DUZ'
A/Cross-references: UNIPROT:P18962; UNIPARC:UPI0000031A5F; EMBL:U10399; NID:g500689; PID
R/Roberts, C.J.; Pohlig, G.; Rothman, J.H.; Stevens, T.H.
J. Cell Biol. 108, 1363-1373, 1999
A/Title: Structure, biosynthesis, and localization of dipeptidyl aminopeptidase B, an in
A/Reference number: A30107; MUID:89174971; PMID:2647766
A/Accession: A30107
A/Molecule type: DNA
A/Residues: 1-82 'H', 84-124, 'N', 126-181, 'LRRLET', 189-199, 'N', 201-365, 'DFKRGKERKE', 376-57
A/Cross-references: UNIPARC:UPI0000172A3C; EMBL:X15484
A/Note: the authors translated the codon ACC for residue 572 as Asn
C/Genetics:
A/Gene: SGD:DAP2; STE13; MIPS:YHR028C
A/Cross-references: SGD:S0001070; MIPS:YHR028C
A/Map position: 8R

C/Superfamily: dipeptidyl-peptidase IV
C/Keywords: dipeptidylpeptide hydrolase; glycoprotein; transmembrane protein; yeast vacu
F:30-45/Domain: transmembrane #status predicted <TM>
F:63,79,110,139,392,421/Binding site: carbohydrate (Asn) (covalent) #status predicted
Alignment Scores:
Pred. No.: 7,41e-21 Length: 818
Score: 310.00 Matches: 98
Percent Similarity: 37.4% Conservative: 38
Best Local Similarity: 26.9% Mismatches: 122
Query Match: 14.5% Indels: 106
DB: 1 Gaps: 13
US-10-825-632-4 (1-1197) x A30107 (1-818)
Qy 3 TTTGAAGGCCAACAGACTCCCTTTAGAGCATCATCCTGTAGTAGTCAGTACGTTAAAT 62
Db 451 PheIleSerThrArgLysSerSerThrGluArgHisValTyrrIleAspLeuArgSer 470
Qy 63 CTGGAGAGGTGACAGGCTGACTGACCGTGGCTACTCA-----CATTCTTGCTGCATC 116
Db 471 ProAsnGluIleIleGluValThrAspThrSerGluAspGlyValTyrrAspValSerPhe 490
Qy 117 AGTCAGCAGCTGCTACTCTTTATAGTAAGTATAGTAACACAGAGAAATCCACAC----- 170
Db 491 SerSerGlyArgArgPheGlyLeuLeuThrTyrrLysGlyProLysValProTyrrGlnLys 510
Qy 171 -----TGT-----GTG 176
Db 511 IleValAspPheHisSerArgLysAlaGluLysCysAspLysGlyAsnValLeuGlyLys 530
Qy 177 TCCCTTTTACAAGCTATCAAGTCTCTGAAGATGATGCCAACCTGTCACAAACAAAGAAATTTGG 236
Db 531 SerLeuTyrrHisLeuGluLysAsnGlu----- 539
Qy 237 GCCACCATTTTGGATTGACAGCGTCTCTCTCTGACTACTATCTCTCCAGAAATTTTCTCT 296
Db 540 -----ValLeuThrLysIle-----LeuGluAspTyrrAlaValProArgLysSerPhe 555
Qy 297 TTTGAAGTACTACTCGATTACATGTTATGG-----ATGCTCTACAAGCCTCATGAT 350
Db 556 ArgGluLeuAsnLeuGlyLysAspGluPheGlyLysAspIleLeuValAsnSerTyrrGlu 575
Qy 351 CTACAGCCT-----GGAAAGAAATATCTCTACTGCTGCTTCATA 389
Db 576 IleLeuProAsnAspPheAspGluThrLeuSerAspHisTyrrProValPhePhePheala 595
Qy 390 TATGGTGGTCTCTCAG----- 404
Db 596 TyrGlyGlyProAsnSerGlnGlnValValLysThrPheSerValGlyPheAsnGluVal 615
Qy 404 ----- 404
Db 616 ValAlaSerGlnLeuAsnAlaIleValValValAspGlyArgGlyThrGlyPheLys 635
Qy 405 -----GGTCAAAATAGAAATTTGACATCAG 428
Db 636 GlyGlnAspPheArgSerLeuValArgAspArgLeuGlyAspTyrrGluAlaArgAspGln 655
Qy 429 GTGAAGGATCCCAATATCTAGCTTCTCGATATGATTTTCATTGATTCGTTAGTCGTGGGC 488
Db 656 IleSerAlaAlaSerLeuTyrrGlySer--LeuThrPheValAspProGlnLysIleSer 674
Qy 489 ATCCACCGCTGGTCTTATGGAGGATACCTCTCCCTGATGGCATTAATATCAGAGGTCAGAT 548
Db 675 LeuPheGlyTrpSerTyrrGlyGlyTyrrLeuThrLeuLysThrLeuGluLysAspGlyGly 694
Qy 549 ---ATCTTCAGGGTGTGCTATTGTCTGGGGCCCGCCAGTCACCTCTGTGATCTTCTATGATACA 605
Db 695 ArgHisPheLysTyrrGlyMetSerValAlaProValThrAspTrpArgPheTyrrAspSer 714
Qy 606 GGATACACGGAACGTTATATGGTTCACCTGACCAAGAAATGAACAGGCTATTACTTAGGA 665

Db 715 ValTyrThrGluArgTyrMetHisThrProGlnGluAsnPheAspGlyTyrValGluSer 734
QY 666 TCCTGTG-----GCCATGCAAGCAGAAAGTTCCCTCTGACCAAAATCGTTTA 713
Db 735 SerValHisAsnValThrAlaLeu-----AlaGlnAlaAsnArgPhe 748
QY 714 CTCTCTTACATGGTTTCCTGGATGAGATGCAATTTTGGACATACAGTATATTA 773
Db 749 LeuLeuMetHisGlyThrGlyAspAsnValHisPheGlnAsnSerLeuLysPheLeu 768
QY 774 AGTTTTTACTAGGCTGGA--AAGCCATATGATTTACAGATCTATCTCTCAGGAGAGA 830
Db 769 AspLeuLeuAspLeuAsnGlyValGluAsnTyrAspValHisValPheProAspSerAsp 788
QY 831 CACAGCATAAGA 842
Db 789 HisSerIleArg 792

RESULT 11
T25173
hypothetical protein T23F1.7a - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T25173
R:Wilkinson, J.
submitted to the EMBL Data Library, October 1996
A:Reference number: Z19990
A:Accession: T25173
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-779 <WIL>
A:Cross-references: UNIPROT:O18119; UNIPARC:UPI000002A220; EMBL:Z81129; PIDN:CAB03411.1;
C:Experimental source: clone T23F1
C:Genetics: T25173
A:Gene: CESP:T23F1.7a
A:Map position: 5
A:Introns: 13/3; 52/1; 111/3; 151/2; 177/3; 346/3; 504/1; 537/1; 628/2; 683/1; 723/1
C:Superfamily: dipeptidyl-peptidase IV

Alignment Scores:
Pred. No.: 3,02e-20 Length: 779
Score: 303.50 Matches: 93
Percent Similarity: 42.8% Conservative: 62
Best Local Similarity: 25.7% Mismatches: 118
Query Match: 14.2% Indels: 89
DB: 2 Gaps: 14

US-10-825-632-4 (1-1197) x T25173 (1-779)
QY 21 TCCCTTTAGAGCATCACCTGTACGTAGTCTAGTACGTAAATCTCGAGAGGTGACAAGG 80
Db 431 AlaProLysProSerHisArgSerLeuPheSerTyr-----Ser 443
QY 81 CTGACTGACCGTGGCTACTCATCTTCTGCTGCATCAGT-----CAGCACTGTGAC 131
Db 444 LeuAlaAspGluSerArgAsnSerAlaTyrCysIleSerCysSerIleLysAsnCysThr 463
QY 132 TTCTTTATAGTAAGTATAGTAACCAAGAAATCCACACTGTGTGCTCCCTTACAGCTA 191
Db 464 TrpAlaGlnAlaGlnMetAspAspGlnMetLysThrAlaIleValSerCysLysGlyPro 483
QY 192 TCNAGTCTCTCAAGAT-----GACCCAACTTGCACAAACAAG 227
Db 484 AlaAlaProHisThrAlaIleValAsnLeuThrArgMetAspSerAspLysLysThrGlu 503
QY 228 -----GAATTTGGGCCCAACATTTTGGATTCAGACAGTCTCTCT 266
Db 504 HisAlaAsnLeuLeuTyrAspLysThrTyrGlnAsnArgValGluGluAlaGly---Leu 522
QY 267 CTGACTATPACTCTCCAGAAATTTCTCTTTTGAAGTACTACTGGATTTCATTTGAT 326
Db 523 ProValIleIleLysGluThrIle-----LysIleSerAspAspPheAlaLeu 539

QY 327 GGGATGCTCTCAAGCCTCATGATCTACAGCTCGAAGAAGATAT-----CCTACT 377
Db 540 IleLysLysSerIleProLysAspIleTyrAsnArgAspLysHisGlnAlaIleProLeu 559
QY 378 GTGCTGTTCTATATATGGTGGTCTCTCAG----- 404
Db 560 IleValHisValTyrGlyGlyProAsnAspGlnAsnThrLysGluAlaThrGlnIleGly 579
QY 404 ----- 404
Db 580 IleGluGluValValAlaSerAlaSerGlnAlaAlaIleLeuArgIleAspGlyArgGly 599
QY 405 -----GGTCAATAGAA 416
Db 600 SerGlyGlyArgGlyTrpLysTyrArgSerAlaIleTyrGlyGlnLeuGlyThrValGlu 619
QY 417 ATTGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATAT---GATTTCATTGAC 473
Db 620 ValGluAspGlnIleLysAlaIleLysValValLeuArgLeuTyrArgHisLeuLeuAsp 639
QY 474 TTAGATCGTGTGGGCATCCACGGCTGGTCTATGAGGATACCTCTCCCTGATGGCATT 533
Db 640 AlaArgArgValAlaValPheGlyTrpSerTyrGlyGlyPheMetThrLeuSerMetVal 659
QY 534 ATGCAGAGGTCAGAT---ATCTTCAGGTTGCTATTGCTGGGGCCCGCCAGTCTCTGG 590
Db 660 AsnGluAlaProGlnPhePheLysCysAlaValSerValAlaProValThrAsnPhe 679
QY 591 ATCTTCTATGATACAGGATACACGAAAGCTTATATGGTGCACCTGACCAAGATGAACAG 650
Db 680 AlaTyrTyrAspAlaThrTyrThrGluArgTyrMetGly-----AspAlaProLeuGlu 697
QY 651 GCGTATTACTTAGATCTGTGCGCCATGCAAGCAAAAGTTCCCTCTCTGAACCAAAATCGT 710
Db 698 SerTyr-----SerAspValThrLysLysLeuAspAsnPheLysSer-----ThrArg 713
QY 711 TTACTGCTCTTACATGGTTTCTCGATGAGAATGTCCATTTTGCACATACAGTATATTA 770
Db 714 LeuLeuLeuMetHisGlyLeuLeuAspAspAsnValHisPheGlnAsnSerAlaIleLeu 733
QY 771 CTGAGTTTTTTAGTGGGCTCGAAAGCCATATGATTTTACAGATCTATCTCTCAGAGAGA 830
Db 734 IleAspGluLeuGlnAsnArgGlyValAspPheAspLeuMetValTyrProAsnGlnAla 753
QY 831 CACAGCATAAAGATTCCTGAATCGGGAGACATATGAACTGCACTCTTTTGCATCACTT 890
Db 754 HisSerLeuSerSerArgThrSer-----HisValValGlyLysMetThrHisPheLeu 771
QY 891 CAAGAA 896
Db 772 ArgGln 773

RESULT 12
T25174
hypothetical protein T23F1.7b - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T25174
R:Wilkinson, J.
submitted to the EMBL Data Library, October 1996
A:Reference number: Z19990
A:Accession: T25174
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-799 <WIL>
A:Cross-references: UNIPROT:O18119; UNIPARC:UPI0000061190; EMBL:Z81129; PIDN:CAB03412.1;
C:Experimental source: clone T23F1
C:Genetics: T25174
A:Gene: CESP:T23F1.7b
A:Map position: 5
A:Introns: 13/3; 52/1; 111/3; 197/3; 366/3; 524/1; 557/1; 648/2; 703/1; 743/1
C:Superfamily: dipeptidyl-peptidase IV

Alignment Score:

Alignment Scores:			
Pred. No.:	3.04e-20	Length:	799
Score:	303.50	Matches:	93
Percent Similarity:	42.8%	Conservative:	62
Best Local Similarity:	25.7%	Mismatches:	118
Query Match:	14.2%	Indels:	89
DB:	2	Gaps:	14

US-10-825-632-4 (1-1197) x T25174 (1-799)

Qy	21	TCCCTTTAGACGATCACCCTGTACGTAGTTCAGTTACGTAAATCTCGGAGAGGGTGACAAGG	80
Db	451	AlaProLysProSerHisArgSerLeuPheSerTyr-----Ser	463
Qy	81	CTCACTCACCGTGGCTACTCATACATTCTGCTGCACAGT-----CAGCACTGTGAC	131
Db	464	LeuAlaAspGluSerArgAsnSerAlaTyrCysIleSerCysSerIleLysAsnCysThr	483
Qy	132	TTCTTTATAAGTAAGTATAGTAACACAGAAGAATCCACACTGCTGTGTCCCTTTACAAGCTA	191
Db	484	TripAlaGlnAlaGlnMetAspAspGlnMetLysThrAlaIleValSerCysLysGlyPro	503
Qy	192	TCGAAGTCTCGAAGAT-----GACCCAACCTGCCAAAACAAG	227
Db	504	AlaAlaProHisThrAlaIleValAsnLeuThrArgMetAspSerAspLysLysThrGlu	523
Qy	228	-----GAAITTTGGGCCACCATTTTGGATTTCAGCAGGTCTCTCT	266
Db	524	HisAlaAsnLeuLeuTyrAspLysThrTyrGlnAsnArgValGluGluAlaGly---Leu	542
Qy	267	CCTGACTATACTCCTCCAGAAATTTCTCTTTGAAAGTAGTACTGGAATTTACATTGTAT	326
Db	543	ProValIleIleLysGluThrIle-----LysIleSerAspAspPheAspAlaLeu	559
Qy	327	GGGATGCTCTACAGCCTCATGATCTACAGCTCGAAAGAAATAT-----CCTACT	377
Db	560	IleLysLeuSerIleProLysAspIleTyrAsnArgAspLysHisGlnAlaIleProLeu	579
Qy	378	GTGCTGTTCATATATGGTCGTCTCTCAG-----	404
Db	580	IleValHisValTyrGlyGlyProAsnAspGlnAsnThrLysGluAlaThrGlnIleGly	599
Qy	404	-----	404
Db	600	IleGluGluValValAlaSerAlaSerGlnAlaAlaIleLeuArgIleAspGlyArgGly	619
Qy	405	-----GGTCAAATAGAA	416
Db	620	SerGlyGlyArgGlyTrpLysTyrArgSerAlaIleTyrGlyGlnLeuGlyThrValGlu	639
Qy	417	ATTGACGATCAGGTGAAGGACTCCAAATATCTAGTCTCTCGATAT---GATTTTCATTGAC	473
Db	640	ValGluAspGlnIleLysAlaIleLysValValLeuArgLeuTyrArgHisLeuLeuAsp	659
Qy	474	TTAGATCGTGTGGGCATCCACGGCTGGTCCTATGTAGGAGTACCTCTCCCTGATGGCATTA	533
Db	660	AlaArgArgValAlaValPheGlyTrpSerTyrGlyGlyPheMetThrLeuSerMetVal	679
Qy	534	ATGCAGAGGTCAAGT---ATCTTCAGGGTGTCTATGCTGGGGCCCAGTCATCTGTGTGG	590
Db	680	AsnGluAlaProGluGlnPhePheLysCysAlaValSerValAlaProValThrAsnPhe	699
Qy	591	ATCTTCTATCATACAGATACACGNAACGTTATATGGGTCTACCTTGACCAGATGAACAG	650
Db	700	AlaTyrTyrAspAlaThrTyrThrGluArgTyrMetGly-----AspAlaProLeuGlu	717
Qy	651	GGCTATTACTTAGGATCTGTGGCCATCAAGCAGAAAAAGTTCCCTCTCGAACCAAAATCGT	710
Db	718	SerTyr-----SerAspValThrLysLysLeuAspAsnPheLysSer-----ThrArg	733
Qy	711	TTACTGCTCTTACATGGTTTTCTCGTAGAATGTCCATTTTGCACATACACAGTATATTA	770
Db	734	LeuLeuLeuMetHisGluLeuLeuAspAsnValHisPheGlnAsnSerAlaIleLeu	753

```
Db 633 ILsAlaTyxAlaArgileAspValArgGlyThrGlyGlyArgGlyTyrAspValValysGlu 652
Qy 405 -----CGTCAATAGAAATTGAGATCAGGTGGAGGACTCCCAATAT 446
Db 653 AlavalTyxArglyLeuGlyAspAlaGluValValAspThrLeuAspMetIleArgAla 672
Qy 447 CTAGCTCTCGATATGATTTCATGACTTAGATCGTGGGCGATCCACGGCTGGTCTAT 506
Db 673 PheIleAsnThrPheGlyPheIleAspGluAspArgIleAlaValMetGlyTyrSerTyr 692
Qy 507 GAGGATACCTC---TCCCTGATGCAATTAATGAGGTCAGATCATCTTCAGGGTTGCT 563
Db 693 GlyGlyPheLeuThrSerLysIleAlaIleLysAspGlnGlyGluLeuValLysCysAla 712
Qy 564 ATTGGTGGGCGCCAGCTACTCTGTGGATCTTCTATGATACAGATACACGGAAACGTTAT 623
Db 713 IleSerIleAlaProValThrAspPheLysTyrTyrAspSerAlaTyrThrGluArgTyr 732
Qy 624 ATGGGTCAACCTGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCA 683
Db 733 LeuGlyGlnProAlaGluAsnLeuGlnGlyTyrIleAsnThrAsnValIleProHisAla 752
Qy 684 GAAAGTCCCTCTGAACCAATCGTTTACTGCTCTTACATGGTTTCTCGGATGAGAT 743
Db 753 ArgAsnValThrAsnVal-----LysTyrLeuLeuAlaHisGlyGluArgAspAsn 770
Qy 744 GTCCATTTGCATACACAGTATATTAAGTGGTCTTACATGGTTTCTCGGATGAGAT 803
Db 771 ValHisTyrGlnAsnSerAlaArgTrpSerGluAlaLeuGlnGlnAsnGlyIleHisPhe 790
Qy 804 GATTACAGATCTATCTCTCAGAGAGACACAGCATAGAGTCTCTGAATCGGAGAACAT 863
Db 791 ThrGlnLeuValTyrAlaAsnGluAlaHisSerLeu-----SerHis 804
Qy 864 TATGAAGTGCATCTTTTGCATCTTTCACCTTCAAGAAACCTT 902
Db 805 LysLeuPheHisLeuTyrGlyGluValGlnArgPheLeu 817

RESULT 14
A41793
dipeptidyl aminopeptidase-like protein (EC 3.4.14.-) short form, DPPX-S - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C:Accession: A41793
R:Wada, K.; Yokotani, N.; Hunter, C.; Doi, K.; Wentholt, R.J.; Shimasaki, S.
Proc. Natl. Acad. Sci. U.S.A. 89, 197-201, 1992
A:Title: Differential expression of two distinct forms of mRNA encoding members of a dip
A:Reference number: A41793; MUID:92108018; PMID:1729689
A:Accession: A41793
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-803 <WAD>
A:Cross-references: UNIPROT:P42659; UNIPARC:UPI000002A83C; GB:M76429; NID:94087119; PIDN:
A:Note: sequence extracted from NCBI Backbone (NCBI:75138)
C:Superfamily: dipeptidyl-peptidase IV
C:Keywords: dipeptidylpeptide hydrolase; glycoprotein
F:257,342/Binding site: carbohydrate (Asn) (covalent) #status predicted

Alignment Scores:
Pred. No.: 2,58e-16 Length: 803
Score: 262.00 Matches: 89
Percent Similarity: 37.6% Conservative: 45
Best Local Similarity: 25.0% Mismatches: 152
Query Match: 12.2% Indels: 70
DB: 2 Gaps: 9

US-10-825-632-4 (1-1197) x A41793 (1-803)
```

```
Qy 3 TTTGAAGGACCAAGACTCCCTTTAGAGCATCACTGAGTACGTAGTACGTAAAT 62
Db 440 PheLeuSerThrGluAspLeuProArgArgGlnLeuTyrSerAlaSerThrVal--- 458
```

```
Qy 63 CCTGGAGGTGACRAGG-----CTGACTGACCGTGGCTACTACAT 104
Db 459 ---GlySerPheAsnArgGlnCysLeuSerCysAspLeuValAspAsnCysThrTyrPhe 477
Qy 105 TCTTGTGCATCAGTCAGACTGTGACTCTTTTATAGTAAGTATAGTAACCAAGAAAT 164
Db 478 SerAlaSerPheSerProGlyAlaAspPhePheLeuLeuLysCysGluGlyProGlyVal 497
Qy 165 CCACACTGTGTGCTCTTTCAAGACTATCAAGTCTCTGAAGATGACCCCAACTTGGCAACA 224
Db 498 Pro---ThrValSerValHisAsnThrThrAspLysLysLysMetPheAspLeuGluThr 516
Qy 225 AAGGAA---TTTTGGGCCACCATTTTGGATTCAGAGGTCTCTCTCTGACTATATCCT 281
Db 517 AsnGluHisValGlnLysAlaIleSerAspArgGlnMetProLysValGluTyrArgLys 536
Qy 282 CCAGAAATTTCTCTTTTGAAGTACTACTGGATTACATTTGATGGATGCTCTTACAAG 341
Db 537 IleGlu-----ThrAspAspTyrAsnLeuProIleGlnIleLeuLys 550
Qy 342 CCTCATGATCTACAGCCTCGAAAGAAATATCTACTGTGCTGTTCATATATGGTGGTCT 401
Db 551 ProAlaThrPheThrAspThrAlaHisTyrProLeuLeuLeuValValAspGlyThrPro 570
Qy 401 -----ThrAspAspTyrAsnLeuProIleGlnIleLeuLys 401
Db 571 GlySerGlnSerValAlaGluLysPheAlaValThrTrpGluThrValMetValSerSer 590
Qy 401 -----CAGGTCMAATAGAAATTGACGATCAGGTGGAGGACTC 440
Db 591 HisGlyAlaValValLysCysAspGlyArgGlySerGlyPheGlnGlyThrArgLeu 610
Qy 402 -----CAGGTCMAATAGAAATTGACGATCAGGTGGAGGACTC 440
Db 611 LeuHisGluValArgArgLeuGlySerLeuGluGluLysAspGlnMetGluAlaVal 630
Qy 441 CAATATCTAGCTTCTCGATATGATTTCATGACTTAGATCGTGTGGGATCCACGGTGG 500
Db 631 ArgValMetLeu---LysGluProTyrIleAspLysThrArgValAlaValPheGlyLys 649
Qy 501 TCCTATGGAGGATACCTCTCCCTGATGGCATTAATGCAGAGGTGCAGAT----- 548
Db 650 AspTyrGlyGlyTyrLeuSerThrTyrLeuLeuProAlaLysGlyAspGlyGlnAlaPro 669
Qy 549 ATCTTCAAGGTTGCTATTGCTGGGCGCCAGTCACCTCTGTGGATCTTCTATGATACAGGA 608
Db 670 ValPheSerCysGlySerAlaLeuSerProIleThrAspPheLysLeuTyrAlaSerAla 689
Qy 609 TACACGGAACCTTATATGGTGCCTGACCAAGATGAACAGGGCTATTACTTAGGATCT 668
Db 690 PheSerGluArgTyrLeuGlyLeuHisGlyLeuAspAsnArgAlaTyrGluMetAlaLys 709
Qy 669 GTGGCCATGCAACGAAAGTTCCCTCTGAAACCAATCGTTTACTCTCTTACATGGT 728
Db 710 ValAlaHisArgValSerAlaLeu-----GluGlyGlnGlnPheLeuValIleHisAla 727
Qy 729 TTCTCGATGCAAGATGCTCCATTTTCACATACCACTATATTACTGAGTTTGTAGTGAGG 788
Db 728 ThrAlaAspGluLysIleHisPheGlnHisThrAlaGluLeuIleThrGlnLeuIleLys 747
Qy 789 CTGGAAAGCCATATGATTACAGATCTATCTCTAGGAGACACACAGCATAGAGTTCCT 848
Db 748 GlyLysAlaAsnTyrSerLeuGlnIleTyrProAspGluSerHisTyrPheSerSerAla 767
Qy 849 GAATCGGGAGAACATTATGAACCTGTCATCTTTTGGCACTACCTCTTCAAGAA 896
Db 768 AlaLeuGlnGlnHisLeuHisArgSerIleLeuGlyPhePheValGlu 783

RESULT 15
I68600
dipeptidyl aminopeptidase like protein - human
C:Species: Homo sapiens (man)
C:Date: 06-Sep-1996 #sequence_revision 06-Sep-1996 #text_change 21-Jul-2000
```

C;Accession: I68600
R;Yokotani, N.; Doi, K.; Wentholt, R.J.; Wada, K.
Hum. Mol. Genet. 2, 1037-1039, 1993
A;Title: Non-conservation of a catalytic residue in a dipeptidyl aminopeptidase IV-related
A;Reference number: I54331; MUID: 93372805; PMID: 8103397
A;Accession: I68600
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-803 <RES>
A;Cross-references: UNIPARC:UPI000016A7P9; GB:M96860; NID:G306707; PIDN:AAA35761.1; PID:
C;Superfamily: dipeptidyl-peptidase IV

Alignment Scores:
Pred. No.: 5,44e-15 Length: 803
Score: 248.00 Matches: 84
Percent Similarity: 37.4% Conservative: 53
Best Local Similarity: 23.0% Mismatches: 139
Query Match: 11.6% Indels: 90
DB: 2 Gaps: 10

US-10-825-632-4 (1-1197) x I68600 (1-803)

QY	3	TTTGAAGGACCAAGACTCCCTTTAGAGCATCACCTGTACGTAGTCAGTTACGTAAT	62
DB	440	PheLeuSerThrGluAspLeuProArgArgGlnLeuTyrSerAlaAsn-----	456
QY	63	CCTGGAGGTGACAAAGGTGACTGACCGTGGCTACTCACAT-----TCTGCTGC	113
DB	457	-----ThrGluGlyAsnPheAsnArgGlnCysLeuSerCysAsp	469
QY	114	ATCAGTCAGCACTGTGACTCTTTTATAGTAGTATAGTACCAAGAAATCCACACTGT	173
DB	470	LeuValGluAsnCysThrTyrPheSerAlaSerPheSerHisSerMetAsp-----	486
QY	174	GTCTCCTTTTACAGCTATCAAGTCTCT-----GAAGATGACCCCACT	215
DB	487	PhePheLeuLeuLysCysGluGlyProGlyValProMetValThrValHisAsnThrThr	506
QY	216	TGCAAAACAAAGAAATTT-----TGGGCCCACTTTTGGAT	251
DB	507	AspLysLysLysMetPheAspLeuGluThrAsnGluHisValLysLysAlaIleAsnAsp	526
QY	252	TCAGCAGGTCTCTCTGACTACTACTCTCCAGAAATTTCTCTTTTGAAAGTACTACT	311
DB	527	ArgGlnMetProLysValGluTyrArgAspIleGluIle-----Asp	540
QY	312	GDATTTACATGTGGATGCTTACAGCCTCATGATCTACAGCCTGGAAGAAATAT	371
DB	541	AspTyrAsnLeuProMetGlnIleLeuLysProAlaThrPheThrAspThrThrHisTyr	560
QY	372	CCTACTGCTGCTTCATATATGCTGCTCCTCAGGTCAA-----	410
DB	561	ProLeuLeuLeuValValAspGlyThrProGlySerGlnSerValAlaGluLysPheGlu	580
QY	410	-----	410
DB	581	ValSerTrpGluThrValMetValSerSerHisGlyAlaValValLysCysAspGly	600
QY	410	-----	410
DB	601	ArgGlySerGlyPheGlnGlyThrLysLeuLeuHisGluValArgArgLeuGlyLeu	620
QY	411	ATAGAAATTGACGATCAGGTGGAAGGACTCCCAATATCTAGCTTCTCGATATGATTTCA	470
DB	621	LeuGluGluLysAspGlnMetGluAlaValArgThrMetLeu---LysGluGlnTyrIle	639
QY	471	GACTTAGATCGTGGGATCCACGGCTGCTTATGAGGATACCTCTCCCTGATGGCA	530
DB	640	AspArgThrArgValAlaValPheGlyLysAspTyrGlyGlyTyrLeuSerThrTyrIle	659
QY	531	TTAATGCAGAGTTCAGAT-----ATCTTTCAGGGTTGCTATTGCTGGGGCCCA	578
DB	660	LeuProAlaLysGlyGluAsnGlnGlyGlnThrPheThrCysGlySerAlaLeuSerPro	679

QY	579	GTCACTCTGTGGATCTTCTATGATACAGGATACCGAACGTTATATGGTCAACCTGAC	638
DB	680	IleThrAspPheLysLeuTyrAlaSerAlaPheSerGluArgTyrLeuGlyLeuHisGly	699
QY	639	CAGAATCAACAGGGCTATTACTTTAGGATCTGTGGCCATGCAAGCAGAAAAAGTTCCCTCT	698
DB	700	LeuAspAsnArgAlaTyrGluMetThrLysValAlaHisArgValSerAlaLeu-----	717
QY	699	GAACCAAAATCGTTTACTGCTCTTTACATGGTTTCTGGATGAGAATGTCATTTTGCACAT	758
DB	718	GluGluGlnGlnPheLeuIleHisProThrAlaAspGluLysIleHisPheGlnHis	737
QY	759	ACCAGTATATTAAGTATTTTACTGAGGGTGGAAAGCCCATATGATTACAGATCTAT	818
DB	738	ThrAlaGluLeuIleThrGlnLeuIleArgGlyLysAlaAsnTyrSerLeuGlnIleTyr	757
QY	819	CCTCAGGAGAGACACAGCATTAAGAGTTCTCTGAATCGGAGAACATTATGAACCTGCATCT	878
DB	758	ProAspGluSerHisTyrPheThrSerSerSerLeuLysGlnHisLeuTyrArgSerIle	777
QY	879	TTGCACCTACCTTCAAGAA	896
DB	778	IleAsnPhePheValGlu	783

Search completed: May 2, 2006, 01:38:07
Job time : 61.6174 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocorelation Ltd.

OM nucleic - protein search, using frame_plus.n2p model

Run on: May 2, 2006, 00:42:23 ; Search time 55.1366 Seconds
(without alignments)
3063.366 Million cell updates/sec

Title: US-10-825-632-4

Perfect score: 2140

Sequence: 1 attttgaaggcaccacaaagac.....ttactataaaaaaaaaaaaaa 1197

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 4332886

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp
-Q=/abs/ABSSWEB_spool/US10825632/runat_01052006_105946_3226/app_query.fasta_1
-DB=Uniprot -QFMT=fastan -SUFFIX=n2p.rup -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bites -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs02p
-USR=US10825632 @CGN 1 1 694 @runat_01052006_105946_3226 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG -DSV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DSLEX=7

Database : Uniprot 05.80.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1645.5	76.9	898	1 DPP8 HUMAN	Q6v1x1 homo sapien
2	1605.5	75.0	892	1 DPP8_MOUSE	Q80ya7 mus musculus
3	1094.5	51.1	862	1 DPP9_MOUSE	Q8bvq4 mus musculus
4	1093.5	51.1	863	1 DPP9_HUMAN	Q8et12 homo sapien
5	1084.5	50.7	923	2 Q4SBM6_TETNG	Q4sbm6 tetraodon n
6	1059	49.5	847	2 Q6GR22_XENLA	Q6gr22 xenopus lae
7	1032.5	48.2	508	2 Q75273_HUMAN	Q75273 homo sapien
8	730.5	34.1	621	2 Q7PTT8_ANOGA	Q7ptt8 anopheles g
9	730.5	34.1	886	2 Q7QBK1_ANOGA	Q7qbki anopheles g
10	710.5	33.2	740	2 Q5TTK8_ANOGA	Q5ttk8 anopheles g
11	704.5	32.9	1053	2 Q9VC20_DROME	Q9vc20 drosophila
12	704.5	32.9	1113	2 Q9VC19_DROME	Q9vc19 drosophila
13	639	29.9	557	2 Q5TXJ2_ANOGA	Q5txj2 anopheles g
14	599	28.0	803	2 Q54U01_DICDI	Q54u01 dictyosteli
15	502	23.5	927	2 Q965K3_CAENOR	Q965k3 caenorhabdi
16	502	23.5	931	2 Q44987_CAENOR	Q44987 caenorhabdi

17	484.5	22.6	738	2	Q9A6E0_CAUCR	Q9a6e0 caulobacter
18	458	21.4	895	2	Q61CU7_CABBR	Q61cu7 caenorhabdi
19	453.5	21.2	736	2	Q8A028_BACTN	Q8a028 bacteroides
20	449	21.0	745	2	Q6F317_9PSED	Q6f317 pseudomonas
21	445.5	20.8	596	2	Q6K880_ORYSA	Q6k880 oryza sativ
22	442.5	20.7	745	2	Q5H5M8_XANOR	Q5h5m8 xanthomonas
23	442	20.7	552	2	Q8GUJ7_ARATH	Q8guj7 arabidopsis
24	442	20.7	746	2	Q9FNF6_ARATH	Q9fnf6 arabidopsis
25	441.5	20.6	757	2	Q8PFD7_XANAC	Q8pfd7 xanthomonas
26	440.5	20.6	766	2	Q4TNP1_9SPHN	Q4tnp1 erythrobact
27	436.5	20.4	763	2	Q8EAB7_SHEON	Q8eab7 shewanella
28	433.5	20.3	751	2	Q4UPD3_XANCP	Q4upd3 xanthomonas
29	433.5	20.3	751	2	Q8P3V8_XANCP	Q8p3v8 xanthomonas
30	432.5	20.2	743	2	Q5QX36_IDILO	Q5qx36 idiomarina
31	427.5	20.0	736	2	Q5LGU5_BACFN	Q5lgus bacteroides
32	427.5	20.0	736	2	Q64XP9_BACFR	Q64xp9 bacteroides
33	424	19.8	850	2	Q6H9E3_9TRYP	Q6h9e3 trypanosoma
34	409	19.1	852	2	Q4Q1H9_LEIMA	Q4qlh9 leishmania
35	400	18.7	720	2	Q5NMW8_ZYMMO	Q5nmw8 zymomonas m
36	398.5	18.6	741	2	P95782_XANMA	P95782 xanthomonas
37	393.5	18.4	732	2	Q7MUW6_PORGI	Q7muw6 porphyromon
38	383	17.9	749	2	Q7NEK8_GLOVI	Q7nek8 gloeobacter
39	379	17.7	711	2	Q47900_FLAME	Q47900 flavobacter
40	377.5	17.6	731	2	Q6L872_PREIN	Q6l872 prevotella
41	376	17.6	723	2	Q31048_PORGI	Q31048 porphyromon
42	376	17.6	723	2	Q66223_PORGI	Q66223 porphyromon
43	375	17.5	778	2	Q51ZP7_MAGRI	Q51zp7 magnaporthe
44	374.5	17.5	730	2	Q93JY4_9BACT	Q93jy4 prevotella
45	372	17.4	739	2	Q5FQY6_GLUOX	Q5fqy6 gluconobact

ALIGNMENTS

RESULT 1
DPP8 HUMAN STANDARD: PRT: 898 AA.
ID DP8_HUMAN STANDARD: Q7Z4E1; Q81WG7; Q8NEM5; Q96JX1; Q9HEM2;
AC Q6V1X1; Q7Z4C8; Q7Z4D3; Q7Z4E1; Q81WG7; Q8NEM5; Q96JX1; Q9HEM2;
AC Q9HEM2; Q9HEM4; Q9HEM5; Q9NXP4;
DT 13-SEP-2005 (Rel. 48, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DT 13-SEP-2005 (Rel. 48, Last sequence update)
DE Dipeptidyl peptidase 8 (EC 3.4.14.5) (dipeptidyl peptidase VIII) (DP8)
DE (Prolyl dipeptidase DPP8) (Dipeptidyl peptidase IV-related protein 1) (DPP8-1).
DE Name=DPP8; Synonyms=DPRP1; ORFNames=MSTP097, MSTP135, MSTP141;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 3), NUCLEOTIDE SEQUENCE [MRNA] OF 334-898 (ISOFORM 4), NUCLEOTIDE SEQUENCE [MRNA] OF 540-898 (ISOFORM 5), NUCLEOTIDE SEQUENCE [MRNA] OF 260-792 (ISOFORM 6), FUNCTION, CATALYTIC ACTIVITY, ENZYME REGULATION, TISSUE SPECIFICITY, INDUCTION, AND SUBCELLULAR LOCATION.
RC TISSUE=Placenta;
RX MEDLINE=20467194; PubMed=11012666;
RA Abbott C.A., Yu D.M.T., Woollatt E., Sutherland G.R., McCaughan G.W., Gorrell M.D.;
RT "Cloning, expression and chromosomal localization of a novel human dipeptidyl peptidase (DPP) IV homolog, DPP8.";
RL Eur. J. Biochem. 267:6140-6150(2000).
[2]
RN NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 3), CATALYTIC ACTIVITY, ENZYME REGULATION, BIOPHYSICO-CHEMICAL PROPERTIES, SUBCELLULAR LOCATION, AND TISSUE SPECIFICITY.
RC TISSUE=retis;
RX PubMed=12662155; DOI=10.1042/BJ20021914;
RA Qi S.Y., Riviere P.J., Trojnar J., Junien J.-L., Akiansanya K.O.;
RT "Cloning and characterization of dipeptidyl peptidase 10, a new member of an emerging subgroup of serine proteases.";
RL Biochem. J. 373:179-189(2003).

[3] NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 1).
RP TISSUE=Testis; Sha J.H., Zhou Z.M., Li J.M.;
RA Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
[4]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 4), AND NUCLEOTIDE
RP SEQUENCE [LARGE SCALE MRNA] OF 211-898 (ISOFORM 2).
RC TISSUE=Hepatoma, and placenta;
EX PubMed=14702033, DOI=10.1038/ngl1285;
RA Oka T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,
RA Nagahori K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M.,
RA Shiratori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H.,
RA Omura Y., Abe K., Kamiyama K., Katsuta N., Sato K., Tanikawa M.,
RA Yamazaki M., Ninomiya K., Ishibashi T., Yamashita H., Murakawa K.,
RA Fujimori K., Tanai H., Kimata M., Watanabe M., Hiraoaka S., Chiba Y.,
RA Iihida S., Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T.,
RA Kusano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,
RA Nomura Y., Togiyasu S., Komai F., Hara R., Takeuchi K., Arita M.,
RA Inose N., Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,
RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohara N., Sano S.,
RA Moriya S., Momiya N., Satoh N., Takami S., Terashima Y., Suzuki O.,
RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
RA Hishigaki H., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
RA Yamazaki M., Watanabe K., Tashiro H., Tanigami A., Fujiwara T.,
RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohnori Y., Okamoto S.,
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikemura Y.,
RA Okitani R., Kawakami T., Noguchi S., Ichi T., Shigeoka K., Senba T.,
RA Tatemura K., Nakajima Y., Mizuno T., Morinaga M., Kawakami B.,
RA Togaishi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
RA Nakaki K., Ikeda I., Nakamura Y., Ohara O., Isogaki S., Sugano S.;
RT "Complete sequencing and characterization of 21,243 full-length human
RT cDNAs.";
RL Nat. Genet. 36:40-45 (2004).
[5]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORMS 1 AND 2).
RC MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RX Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins B.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whitling M., Helton E., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Trinchard J., Schmutz J., Myers R.M.,
RA Butterfield V.S.N., Krzywicki M.J., Skalska U., Smalus D.E.,
RA Scherch A., Schin J.E., Jones S.J.M., Mirra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
[6]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 561-898.
RC TISSUE=Aorta;
RA Zhao B., Xu H.S., Tong Y.K., Sheng H., Qin B.M., Liu Y.Q., Liu B.,
RA Wang X.Y., Zhang Q., Song L.Z., Gao Y., Zhang C.L., Ye J., Ji X.J.,
RA Liu B.H., Lu H., Chen J.Z., Cai M.Q., Zheng W.Y., Teng C.Y., Liu Q.,
RA Yu L.T., Lin J., Gong Q., Zhang A.M., Gao R.L., Hui R.T.;
RT Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.

[7] MUTAGENESIS OF GLU-275; SER-755; ASP-833 AND HIS-865, CATALYTIC
RP ACTIVITY, BIOPHYSICOCHEMICAL PROPERTIES, AND SUBCELLULAR LOCATION.
RX PubMed=12534281; DOI=10.1021/bi026846s;
RA Ajami K., Abbott C.A., Obradovic M., Gysbers V., Kaehne T.,
RA McCaughan G.W., Gorrell M.D.;
RT "Structural requirements for catalysis, expression, and dimerization
RT in the CD26/DPPIV gene family.";
RL Biochemistry 42:694-701 (2003).
[8]
RP BIOPHYSICOCHEMICAL PROPERTIES.
RX PubMed=15039077; DOI=10.1016/j.pep.2003.12.019;
RA Chen Y.-S., Chien C.-H., Goparaju C.M., Hsu J.T.-A., Liang P.-H.,
RA Chen X.;
RT "Purification and characterization of human prolyl dipeptidase DPP8 in
RT Sf9 insect cells.";
RL Protein Expr. Purif. 35:142-146 (2004).
[9]
RP ENZYME REGULATION.
RX PubMed=15664838; DOI=10.1016/j.bmol.2004.11.023;
RA Jiaang W.-T., Chen Y.-S., Hau T., Wu S.-H., Chien C.-H., Chang C.-N.,
RA Chang S.-P., Lee S.-J., Chen X.;
RT "Novel isoindoline compounds for potent and selective inhibition of
RT prolyl dipeptidase DPP8.";
RL Bioorg. Med. Lett. 15:687-691 (2005).
CC -!- FUNCTION: Dipeptidyl peptidase that cleaves off N-terminal
CC dipeptides from proteins having a Pro or Ala residue at position
CC 2. May play a role in T-cell activation and immune function.
CC -!- CATALYTIC ACTIVITY: Release of an N-terminal dipeptide, Xaa-Yaa-|-
CC Zaa-, from a polypeptide, preferentially when Yaa is Pro, provided
CC Zaa is neither Pro nor hydroxyproline.
CC -!- ENZYME REGULATION: Inhibited by zinc. Inhibited by the serine
CC proteinase inhibitor 4-(2-aminoethyl)benzenesulphonyl fluoride
CC (AEBSF), and by di-isopropylfluorophosphate. Specifically inhibited
CC by isoindoline derivatives.
CC -!- BIOPHYSICOCHEMICAL PROPERTIES:
CC Kinetic parameters:
CC KM=208 uM for Ala-Pro-AMC;
CC KM=130 uM for Ala-Pro-AFC;
CC KM=120 uM for H-Ala-Pro-pNa;
CC KM=1420 uM for H-Ala-Ala-pNa;
CC KM=310 uM for H-Arg-Pro-pNa;
CC KM=2050 uM for H-Asp-Pro-pNa;
CC KM=480 uM for H-Gly-Pro-pNa;
CC pH dependence:
CC Optimum pH is 7.4-8.5. Little activity below pH 6.5;
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=6;
CC Name=1:
CC IsoId=Q6V1X1-1; Sequence=Displayed;
CC Name=2:
CC IsoId=Q6V1X1-2; Sequence=VSP_013864;
CC Name=3:
CC IsoId=Q6V1X1-3; Sequence=VSP_013860;
CC Name=4:
CC IsoId=Q6V1X1-4; Sequence=VSP_013862;
CC Name=5:
CC IsoId=Q6V1X1-5; Sequence=VSP_013863;
CC Name=6:
CC IsoId=Q6V1X1-6; Sequence=VSP_013861;
CC -!- TISSUE SPECIFICITY: Ubiquitously expressed, with highest levels in
CC testis, placenta, prostate, muscle and brain.
CC -!- INDUCTION: In activated T-cells.
CC -!- SIMILARITY: Belongs to the peptidase S9B family. DPPIV subfamily.
CC -!- CAUTION: Ref.4 (BAA91059) sequence differs from that shown due to
CC a frameshift in position 486.
CC -!- CAUTION: Ref.6 (AAQ13650 and AAQ13623) sequences differ from that
CC shown due to several frameshifts.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

use as long as its content is in no way modified and this statement is not removed.

CC
CC
DR EMBL; AF221634; AAG29766.1; -; mRNA.
DR EMBL; AF221635; AAG29767.1; -; mRNA.
DR EMBL; AF221636; AAG29768.1; -; mRNA.
DR EMBL; AF221637; AAG29769.1; -; mRNA.
DR EMBL; AY172659; AAO17261.1; -; mRNA.
DR EMBL; AY354202; AAO3887.1; -; mRNA.
DR EMBL; AK000290; BAA91059.1; -; mRNA.
DR EMBL; AK027826; BAB55395.1; ALT_FRAME; mRNA.
DR EMBL; BC030688; BAH30688.3; -; mRNA.
DR EMBL; BC040203; AAH40203.1; ALT_INIT; mRNA.
DR EMBL; AF176779; AAQ13657.1; ALT_INIT; mRNA.
DR EMBL; AF175225; AAQ13650.1; ALT_FRAME; mRNA.
DR EMBL; AF173382; AAQ13623.1; ALT_FRAME; mRNA.

Alignment Scores:
Pred. No.: 6,46e-144 Length: 898
Score: 1645.50 Matches: 310
Percent Similarity: 86.4% Conservative: 0
Best Local Similarity: 86.4% Mismatches: 0
Query Match: 76.9% Indels: 49
DB: 1 Gaps: 1

US-10-825-632-4 (1-1197) x DPP8_HUMAN (1-898)

QY 3 TTTGAAGCCACCAAGACTCCCTTTAGAGCATCCTCTAGCTAGTACGTTACGTAAAT 62
DB PheGluGluThrLysAspSerProLeuGluHisLeuTyrValValSerTyrValAsn 559
QY 63 CTGGAGAGGTGACAAAGCGTCACTGACCGTGGCTACTACATCTTGTGTCAGTCAG 122
DB ProGluGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCysLeuSerGln 579
QY 123 CACTGTGACTCTTTTATAGTATAGTAAACCAAGAAATCCACATCTGTGTCCCTT 182
DB HisCysAspPhePheLeuSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeu 599
QY 183 TACAGCTATCAAGTCTGAGATGACCCCACTTGCAACAAAGAAATTTTGGCCACC 242
DB TyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPheTrpAlaThr 619
QY 243 ATTTGGATTCCAGCAGCTCTCTCTGCTACTACTCTCCAGAAATTTCTCTTTGAA 302
DB IleLeuAspSerAlaGlyProLeuProAspTyrThrProGluLeuPheSerPheGlu 639
QY 303 AATCTACTGATTTAATTGATGATGCTCTCAAGCCTCATGATCTACAGCTGGA 362
DB SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly 659
QY 363 AAGAAATATCTACTGCTGCTTCATATATATGTTGCTCTCAG----- 404
DB LysLysTyrProThrValLeuPheLeuTyrGlyGlyProGlnValGlnLeuValAsn 679
QY 404 ----- 404
DB ArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrValVal 699
QY 404 ----- 404
DB ValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLys 719
QY 405 -----GTCAAATAGAAATTCAGCATCAGTGGAGGACTCCATATCTAGCTTCT 455
DB TyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyrLeuAlaSer 739
QY 456 CGATATGATTTTCATTGACTTATAGTTCGTGGGCATCCAGCGTGGTCTTATGAGGATAC 515
DB ArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrSerTyrGlyGlyTyr 759
QY 516 CTCTCCCTGATGGCATTAATGACAGTCAGATATCTTCAGGGTTGCTATTGCTGGGCCC 575

DB LeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaAlaGlyAla 779
QY 576 CCAGTCACTCTGTGGATCTTTCTATGATACAGGATACACGAACTTATATGGTCCACCT 635
DB ProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMetGlyHisPro 799
QY 636 GACAGAAATGAACAGGGCTATTACTTAGGATCTGTGGCCATCAAGCAGAAAAGTTCCCC 695
DB AspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlyLysPhePro 819
QY 696 TCTGAACCAATCGTTTACTGCTCTTACATGTTCTCTGGATGAGAATGTCATTTCGCA 755
DB SerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnValHisPheAla 839
QY 756 CATACCACTATATTACTAGTGTGTTTGTAGTGGCGTGGAAAGCCATATGATTTACAGATC 815
DB HisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAspLeuGlnIle 859
QY 816 TATCTCAGAGAGACACAGCATTAAGATGTTCTGTAATCGGAGAACATTAAGACTGCAT 875
DB TyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyrGluLeuHis 879
QY 876 CTTTGTGCACTACTCTCAAGAAACCTTGGATCAGCTATTGCTGCTCTAAAAGTGATA 932
DB LeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeuLysValIle 898
RESULT 2
DPP8_MOUSE
ID DPP8_MOUSE STANDARD; PRT; 892 AA.
AC Q80FA7; Q9D4G6;
DT 13-SEP-2005 (Rel. 48, Created)
DT 13-SEP-2005 (Rel. 48, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Dipeptidyl peptidase 8 (EC 3.4.14.5) (Dipeptidyl peptidase VIII)
DE (DPS)
GN Name=Dpp8;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridea; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA Balderelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.W.,
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmond S., Guetincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasawa Y., Kedierski R.M., King B.L.,
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sadelin A., Schneider C., Semple C.A.I., Setou M., Shimada K.,
RA Sulcanea R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."

RL Nature 420:563-573 (2002).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC STRAIN=C57BL/6; TISSUE=Brain;
RX MDLINE=22368257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A., Rubin G.M., Hong L.,
RA Scapleton M., Soares M.B., Bonaldo A.F., Casavant T.L., Scheetz T.E.,
RA Brown S., Loquellan T.B., Toshynski S., Carninci P., Prange C.,
RA Rana S.S., Loughran J.A., Peters G.J., Abramson R.D., Malleby S.J.,
RA Richards S., McEwen P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Villalón D.K., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Fahney J.C., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Whiting J., Heiton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
[3]
RX PHOSPHORYLATION SITE TYR-325.
RX PubMed=14729942; DOI=10.1074/mcp.D300003-MCP200;
RX Shu H., Chen S., Bi Q., Mumby M., Brekken D.L.;
RT "Identification of phosphoproteins and their phosphorylation sites in
RT the WEHI-231 B lymphoma cell line.";
RL Mol. Cell. Proteomics 3:279-286 (2004).
CC -1- FUNCTION: Dipeptidyl peptidase that cleaves off N-terminal
CC dipeptides from proteins having a Pro or Ala residue at position
CC 2. May play a role in T-cell activation and immune function (By
CC similarity).
CC -1- CATALYTIC ACTIVITY: Release of an N-terminal dipeptide, Xaa-Yaa-|-
CC Zaa- from a polypeptide, preferentially when Yaa is Pro, provided
CC Zaa is neither Pro nor hydroxyproline.
CC -1- ENZYME REGULATION: Inhibited by zinc. Inhibited by the serine
CC proteinase inhibitor 4-(2-aminoethyl)benzenesulphonyl fluoride
CC (AEBSP) and by di-isopropyl fluorophosphate. Specifically inhibited
CC by isocitriding derivatives (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the peptidase S98 family. DPPIV subfamily.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
DR EMBL; AK016546; BAB30295.2; -; mRNA.
DR EMBL; BC043124; AAH43124.1; -; mRNA.
DR EMBL; BC059222; AAH59222.1; -; mRNA.
DR MEROPS; S09.018; -.
DR Ensembl; ENSMUSG00000032393; Mus musculus.
DR MGI; MGI:1921638; Dpp8.
DR InterPro; IPR001375; Peptidase_S9.
DR InterPro; IPR002469; Peptidase_S98.
DR InterPro; IPR000379; Ser_estrs.
DR Pfam; PF00930; DPPIV_N; 1.
DR Pfam; PF00326; Peptidase_S9; 1.
KW Aminopeptidase; Hydrolase; Phosphorylation; Protease; Serine protease.
FT ACT SITE 749 749 Charge relay system (By similarity).
FT ACT SITE 827 827 Charge relay system (By similarity).
FT ACT SITE 859 859 Charge relay system (By similarity).
FT MOD_RES 325 325 Phosphotyrosine.
FT CONFLICT 87 87 G -> R (in Ref. 1).
SQ SEQUENCE 892 AA; 102186 MW; 59081CD9792E03ED CRC64;

Alignment Scores:

Pred. No.: 3.42e-140 Length: 892

Score: 1605.50 Matches: 302
Percent Similarity: 85.0% Conservative: 3
Best Local Similarity: 84.1% Mismatches: 5
Query Match: 75.0% Indels: 49
DB: 1 Gaps: 1
US-10-825-632-4 (1-1197) x DPP8_MOUSE (1-892)
Qy 3 TTTGAAGGACCAAGAGATCCCTTTAGAGCATCACCTGTAGTACGTAGTACGTAGTAAAT 62
Db 534 PheGluGlyThrLysAspSerProLeuGluHisLeuTyrValThrSerTyrAlaAsn 553
Qy 63 CTGGAGAGGTCACAAAGGTCGACCGTGGCTACTACATCTTCTGCTCATCAGTCAG 122
Db 554 ProGlyGluValValArgLeuThrAspArgGlyTyrSerHisSerCysCysLeuSerArg 573
Qy 123 CACTGTGACTCTTTTATAAGTAAGTATAGTAACAGAGAAGATCCACACTCTGTGTCCCTT 182
Db 574 HisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeu 593
Qy 183 TACAAGCTATCAAGTCCTCTGAGATGACCCAACTTGCAGAAACAAAGGAATTTTGGGCCACC 242
Db 594 TyrLysLeuSerSerProGluAspAspProValHisLysThrLysGluPheTyrAlaThr 613
Qy 243 ATTTTGGATTCAGCAGGTCCTCTTCTGACTACTACTCTCCAGAAATTTCTCTTTTGA 302
Db 614 IleLeuAspSerAlaGlyProLeuProAspGlyTyrThrProGluLysLysPheSerPheGlu 633
Qy 303 AGTACTACTGGATTTACATGTTGATGGGATGCTCTACAGCCTCATGATCTACAGCTTGA 362
Db 634 SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly 653
Qy 363 AAGAAATATCTACTGCTGCTTCATATATATGTTGGTCTCTCAG----- 404
Db 654 LysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeuValAsnAsn 673
Qy 404 ----- 404
Db 674 ArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrValVal 693
Qy 404 ----- 404
Db 694 ValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLys 713
Qy 405 -----GGTCAATAGAAATTCAGCATCAGGTGGAGGACCTCCAATATCTAGCTTCT 455
Db 714 TyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyrLeuAlaSer 733
Qy 456 CGATATGATTTCATTCGACTTAGATCGTGTGGGCATCCACGGCTGGTCTTATGGAGGATAC 515
Db 734 GlnTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrSerTyrGlyTyr 753
Qy 516 CTCTCCCTGATGGCATTAAATGAGAGGTCAGATATCTTCAGGTTGCTATTGCTGGGGCC 575
Db 754 LeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAlaGlyAla 773
Qy 576 CCAGTCATCTGTGGATCTTCTATGATACAGGATACACGAGACGTTATATGGTCCACCT 635
Db 774 ProValThrLeuThrPheThrAspThrGlyTyrThrGluArgTyrMetGlyHisPro 793
Qy 636 GACCAGATGACAGGGCTATTACTAGTACTGTGGCATCGAGCGATCGAAGAGAACTTCCCC 695
Db 794 AspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGluLysPhePro 813
Qy 696 TCTGAACCAAAATCGTTTACTGCTCTTACATGTTTCTCGATGAGATGCTCATTTTGA 755
Db 814 SerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnValHisPheAla 833
Qy 756 CATACCAAGTATATTACTAGTGTGTTTGTAGTGAGGGCTGGAAAGCCATATGATTTACAGATC 815
Db 834 HisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAspLeuGlnIle 853
Qy 816 TATCCTCAGGAGAGACACAGCATAGAGTTCTCTGAATCGGGAGAACATTATGAACCTGCAT 875

Db 854 TyrProGlnGluArgHisSerIleArgValProGlnSerGlyClnHisTyrGlnLeuHis 873
QY 876 CTTTTCGCACTACCTTCAAGAAACCTTGGATACGATTGCTGCTCTAAAGTGATA 932
Db 874 LeuLeuHisTyrLeuGlnGlnAsnLeuGlySerArgIleAlaAlaLeuHisValIle 892
RESULT 3
DPP9 MOUSE
ID DPP9 MOUSE STANDARD; PRT; 862 AA.
AC Q8BVG4; Q8BWT9;
DT 13-SEP-2005 (Rel. 48, Created)
DT 13-SEP-2005 (Rel. 48, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Dipeptidyl peptidase 9 (EC 3.4.14.5) (Dipeptidyl peptidase IX) (DP9)
DE (Dipeptidyl peptidase-like protein 9) (DPLP9).
GN Name=Dpp9;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridea; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 1).
RC STRAIN=C57BL/6J; TISSUE=Liver, and Olfactory bulb;
RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nikaide I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Brusci V., Chothia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Giesi C., Godzik A., Gough J.,
RA Grinnond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kurochi H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Kongsaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Nunata K., Okido T., Pavan W.J., Perlea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Sempie C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Wahlstedt C., Wang Y., Watanabe Y., Wells C.,
RA Wu L., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shingawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573 (2002).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 2).
RC TISSUE=Thymus;
RX PubMed=15449545;
RA Okazaki N., Kikuno R., Ohara R., Inamoto S., Koseki H., Hiraoka S.,
RA Saga Y., Kitamura H., Nakagawa T., Nagase T., Ohara O., Koga H.;
RT "Prediction of the coding sequences of mouse homologues of FLJ genes:
RT the complete nucleotide sequences of 110 mouse FLJ-homologous cDNAs
RT identified by screening of terminal sequences of cDNA clones randomly
RT sampled from size-fractionated libraries."
RL DNA Res. 11:127-135 (2004).
RN [3]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 1).
RC STRAIN=FVB/N; TISSUE=Salivary gland;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.P., Jordan H., Moore T., Ma X.S.I., Wang J., Hsieh F.,
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
CC -!- FUNCTION: Dipeptidyl peptidase that cleaves off N-terminal
CC dipeptides from proteins having a Pro or Ala residue at position 2
CC (By similarity).
CC -!- CATALYTIC ACTIVITY: Release of an N-terminal dipeptide, Xaa-Yaa-|-
CC Zaa-, from a polypeptide, preferentially when Yaa is Pro, provided
CC Zaa is neither Pro nor hydroxyproline.
CC -!- ENZYME REGULATION: Inhibited by the serine proteinase inhibitor 4-
CC (2-aminoethyl)benzenesulphonyl fluoride (AEBSPF), and by di-
CC isopropylfluorophosphate (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytosolic (By similarity).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=Q8BVG4-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q8BVG4-2; Sequence=VSP_013870, VSP_013871, VSP_013872;
CC -!- SIMILARITY: Belongs to the peptidase S9B family. DPPIV subfamily.
CC -!- CAUTION: Ref.2 sequence differs from that shown due to a
CC frameshift in position 745.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL: AK050021; BAC34034.1; -; mRNA.
DR EMBL: AK078301; BAC37211.1; -; mRNA.
DR EMBL: AK131178; BAC21428.1; ALT SEQ; mRNA.
DR EMBL: BC057631; AAH57631.1; -; mRNA.
DR MEROPS; S09.019; -; Mus musculus.
CC Ensemble; ENSMUSG00000001229; Mus musculus.
DR MGI; MGI:2443967; Dpp9.
DR InterPro; IPR001375; Peptidase_S9B.
DR InterPro; IPR002469; Peptidase_S9B.
DR InterPro; IPR000379; Ser estrs.
DR Pfam; PF00930; DPPIV_N; I.
DR Pfam; PF00326; Peptidase_S9; 1.
DR Alternative splicing; Amino peptidase; Hydrolase; Protease;
KW Serine protease.
FT ACT_SITE 729 729 Charge relay system (By similarity).
FT ACT_SITE 807 807 Charge relay system (By similarity).
FT ACT_SITE 839 839 Charge relay system (By similarity).
FT VARSPPLIC 1 75 Missing (in isoform 2).
FT FTId=VSP_013870.
FT VARSPPLIC 748 787 VAAGAPVTVMAYDTGTYRMYDVPENNQGYEAGSVAL
FT -> PPHEARSPSLPATTDPRAWASASSSSWWEAKPGTASEG
FT QR (in isoform 2).
FT FTId=VSP_013871.
FT VARSPPLIC 788 862 Missing (in isoform 2).
FT FTId=VSP_013872.
FT CONFLICT 369 369 D -> Y (in Ref. 1; BAC37211).
FT CONFLICT 546 546 S -> Y (in Ref. 2).
FT CONFLICT 777 777 Q -> K (in Ref. 1; BAC37211).
FT SEQUENCE 862 AA; 98001 MW; B1D565E824A834E8 CRC64;

Alignment Scores:

Pred. No.: 1.3e-92 Length: 862
Score: 1094.50 Matches: 200
Percent Similarity: 69.3% Conservative: 42
Best Local Similarity: 57.3% Mismatches: 58
Query Match: 51.1% Indels: 49
DB: 1 Gaps: 1

US-10-825-632-4 (1-1197) x DPP9_MOUSE (1-862)

QY	3	TTTGAAGGACCAAGACCTCCCTTTAGAGATCACCTGTAGTCAGTACGTTACGTAAT	62
DB	514	PheGlnGlyThrIysAspThrProLeuGluHisHisLeuTyrValValSerTyrGluSer	533
QY	63	CTGGAGAGTGACAGGCTGACTGACCTGCTACACATCTTCTGCTCATCAGTCAG	122
DB	534	AlaGlyGluIleValArgLeuThrThrLeuGlyPheSerHisSerCysSerMetSerGln	553
QY	123	CACGTGACTCTTTTATAGTATAGTAAACGAGAGATCCACACTGTGTCCCTT	182
DB	554	SerPheAspMetPheValSerHisTyrSerSerValSerThrProCysValHisVal	573
QY	183	TACAAGTATCAAGTCTCTGAAGATGACCCCACTTCCAAAACAAAGGAATTTTGGCCACC	242
DB	574	TyrlsLeuSerGlyProAspAspProLeuHisLysGlnProArgPheTyrPalaser	593
QY	243	ATTGTGATTACAGAGCTCTCTCTGACTACTCTCTCAGAAATTTCTCTTTGAA	302
DB	594	MetMetGluAlaAlaAsnCysProProAspTyrValProGluIlePheHisPheHis	613
QY	303	AGTACTACTGGATTTACATTGTATGGATGCTCTCAAGCCTCATGATCTACAGCTGGA	362
DB	614	ThrArgAlaAspValGlnLeuTyrGlyMetIleTyrLysProHisThrLeuGlnProGly	633
QY	363	AAGAAATATCTACTGTCTGTTCATATATATGTGTCTCTCAG-----	404
DB	634	ArgLysHisProThrValLeuPheValTyrGlyGlyProGlnValGlnLeuValAsnAsn	653
QY	404	-----	404
DB	654	SerPheLysGlyIleLysTyrLeuArgLeuAsnThrLeuAlaSerLeuGlyTyrAlaVal	673
QY	404	-----	404
DB	674	ValValIleAspGlyArgGlySerCysGlnArgGlyLeuHisPheGluGlyAlaLeuLys	693
QY	405	-----GGTCAATAGAAATGACGATCAGGTGGAGGACTCCATATCTAGCTTCT	455
DB	694	AsnGlnMetGlyGlnValGluIleGluAspGlnValGluGlyLeuGlnTyrValAlaGlu	713
QY	456	CGATATGATTCATTGACTTAGATCGTCTGGCATCCAGCGCTGCTCTATGGAGGATAC	515
DB	714	LysTyrGlyPheIleAspLeuSerArgValAlaIleHisGlyIrpSerTyrGlyLysPhe	733
QY	516	CTCTCCCTGATGGCAATTAATGCAGAGTCAGATATCTCAGGTTGCTATTGCTGGGCC	575
DB	734	LeuSerLeuMetGlyLeuIleHisLysProGlnValPheLysValAlaIleAlaGlyAla	753
QY	576	CCAGTCACTCTGTGGATCTTCTATATGATACAGGATACAGGAACGTTATATGGGTCCACCT	635
DB	754	ProValThrValTyrMetAlaTyrAspThrGlyTyrThrGluArgTyrMetAspValPro	773
QY	636	GACCAGATGAACAGGCTATTACTTAGGATCTGTGGCCATGCACAGCAGAAAGTCCCC	695
DB	774	GluAsnGlnGlnGlyTyrGluAlaGlySerValAlaLeuHisValGluLysLeuPro	793
QY	696	TCTGAACCAATCTTTTACTGCTCTTACATGTTCTCTGGATGAGAATGTCCATTTTGA	755
DB	794	AsnGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnValHisPhePhe	813
QY	756	CATACCACTATATTACTAGCTTTTATAGTGGGCTGGAAACCCATATGATTTACAGATC	815
DB	814	HisThrAsnPheLeuValSerGlnLeuIleArgAlaGlyLysProTyrGlnLeuGlnIle	833

QY	816	TATCTCTCAGAGAGACACAGCATATAGACTTCTGAATCGGAGAGACATTTATGACTGCAT	875
DB	834	TyrProAsnGluArgHisSerIleArgCysArgGluSerGlyGluHisTyrGluValThr	853
QY	876	CTTTTGCACCTTACCTTCAAGAAACCTT	902
DB	854	LeuLeuHisPheLeuGlnGluHisLeu	862
RESULT 4			
ID	DPP9_HUMAN	STANDARD;	PRT; 863 AA.
AC	Q86T12; Q6A137; Q6VAL0; Q6ZNT2; Q6ZNU5; Q8N2J7; Q8N3F5; Q8WDX8;		
AC	Q6NT8; Q9BVR3;		
DT	13-SEP-2005 (Rel. 48, Created)		
DT	13-SEP-2005 (Rel. 48, Last annotation update)		
DT	13-SEP-2005 (Rel. 48, Last annotation update)		
DE	Dipeptidyl peptidase 9 (EC 3.4.14.5) (Dipeptidyl peptidase IX) (DPP9)		
DE	Dipeptidyl peptidase 9 (EC 3.4.14.5) (Dipeptidyl peptidase IX) (DPP9)		
DE	Dipeptidyl peptidase 9 (EC 3.4.14.5) (Dipeptidyl peptidase IX) (DPP9)		
DE	IV-related protein 2 (DPRP-2)		
OS	Human sapiens (Human)		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;		
OC	Homo		
OX	NCBI_TaxID=9606;		
RN	[1] NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 1), AND TISSUE SPECIFICITY.		
RP	MEDLINE=2347328; PubMed=12459266; DOI=10.1016/S0378-1119(02)01059-4;		
RA	Olsen C., Wegmann N.		
RT	"Identification and characterization of human DPP9, a novel homologue		
RT	of dipeptidyl peptidase IV."		
RL	Gene 299:185-193 (2002).		
RN	[2]		
RP	NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 1), CATALYTIC ACTIVITY.		
RP	BIOPHYSICO-CHEMICAL PROPERTIES, ENZYME REGULATION, TISSUE SPECIFICITY,		
RP	AND SUBCELLULAR LOCATION.		
RC	TISSUE=Colon;		
RX	PubMed=12662155; DOI=10.1042/Bj20021914;		
RA	Qi S.Y., Riviere P.J., Trojnar J., Junien J.-L., Akimanya K.O.;		
RT	"Cloning and characterization of dipeptidyl peptidase 10, a new member		
RT	of an emerging subgroup of serine proteases."		
RL	Biochem. J. 373:179-189 (2003).		
RN	[3]		
RP	NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 1), PARTIAL NUCLEOTIDE SEQUENCE		
RP	[MRNA] (ISOFORM 2), CATALYTIC ACTIVITY, BIOPHYSICO-CHEMICAL PROPERTIES,		
RP	TISSUE SPECIFICITY, AND SUBCELLULAR LOCATION.		
RX	PubMed=1245913; DOI=10.1016/j.bbaexp.2004.03.010;		
RA	Ajani K., Abbott C.A., McCaughan G.W., Gorrell M.D.;		
RT	"Dipeptidyl peptidase 9 has two forms, a broad tissue distribution,		
RT	cytoplasmic localization and DPPIV-like peptidase activity."		
RL	Biochim. Biophys. Acta 1679:18-28 (2004).		
RN	[4]		
RP	NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 1).		
RC	TISSUE=Placenta; and Skin;		
RX	MEDLINE=23388257; PubMed=12477932; DOI=10.1073/pnas.242603899;		
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,		
RA	Klausner R.D., Collins B.S., Wagner L., Shenmen C.M., Schuler G.D.,		
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,		
RA	Hopkins R.F., Jordan R., Moore T., Max S.I., Wang J., Hsieh F.,		
RA	Diatchenko L., Mariani K., Farmer A.A., Rubin G.M., Hong L.,		
RA	Stapleton M., Soares W.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,		
RA	Brownstein M.J., Udwin T.B., Toshyuk S., Carninci P., Prange C.,		
RA	Rosa S.S., Loquellano N.A., Peters C.J., Abramson R.D., Mullaly S.J.,		
RA	Khas S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,		
RA	Richardson D.K., Muzny K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,		
RA	Villalón D.K., Worley J.C., Sodergren E.J., Lu X., Gibbs R.A.,		
RA	Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,		
RA	Whiting R.W., Touchman J.W., Green E.D., Bouffard G.G.,		
RA	Blakesley A.C., Grimwood J., Schmutz J., Myers R.M., Smalls D.E.,		
RA	Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,		
RA	Schmerch A., Schein J.E., Jones S.J.M., Marra M.A.;		


```
Db 515 PheGlnGlyThrLysAspThrProLeuGluHisLeuTyrValValSerTyrGluAla 534
Qy 63 CTGGAGAGTGCAAGGCTGACTGACCGTGGCTACTACATTTCTTCTGCATCAGTCAG 122
Db 535 AlaGlyGluIleValArgLeuThrThrProGlyPheSerHisSerCysSerMetSerGln 554
Qy 123 CACTGTGACTCTCTTTATAAGTATAGTAACACGAGAAGATCCACACTGTGTCCCTT 182
Db 555 AsnPheAspMetPheValSerHisTyrSerSerValSerThrProCysValHisVal 574
Qy 183 TACAAGGTATCAAGTCTCTCAAGATGACCAACTGCGAAACAAAGGAATTTGGGCCACC 242
Db 575 TyrLysLeuSerGlyProAspAspProLeuHisLysGlnProArgPheThrAlaSer 594
Qy 243 ATTTTGGATTACAGCGTCTCTCTGACTATATCTCTCCAGAAATTTCTCTTTTGA 302
Db 595 MetMetGluAlaAlaSerCysProProAspTyrValProGluIlePheHisPheHis 614
Qy 303 AGTACTACTGGATTTTACATTTGATGGATGCTCTACAGCTCATGATCTACAGCTTGA 362
Db 615 ThrArgSerAspValArgLeuTyrGlyMetIleTyrLysProHisAlaLeuGlnProGly 634
Qy 363 AAGAAATATCTACTGCTGCTTCATATATATGCTGCTCTCAG----- 404
Db 635 LysHisProThrValLeuPheValTyrGlyGlyProGlnValGlnLeuValAsnAsn 654
Qy 404 ----- 404
Db 655 SerPheLysGlyIleLysTyrLeuArgLeuAsnThrLeuAlaSerLeuGlyTyrAlaVal 674
Qy 404 ----- 404
Db 675 ValValIleAspGlyArgGlySerCysGlnArgGlyLeuArgPheGluGlyAlaLeuLys 694
Qy 405 -----GGTCAATAGAAATGACGATCAGGTGGAGGACTCCAATATCTAGCTTCT 455
Db 695 AsnGlnMetGlyGlnValGluIleGluAspGlnValGluGlyLeuGlnPheValAlaGlu 714
Qy 456 CGATATGATTTCATTGACTTACATCGTGTGGGCATCCACGGCTGCTCTATGGAGGATAC 515
Db 715 LysTyrGlyPheIleAspLeuSerArgValAlaIleHisGlyTyrSerTyrGlyGlyPhe 734
Qy 516 CTCTCCTGATGCAATTAATGACAGAGTCAGATCTTTCAGGGTTGCTATTGCTGGGCC 575
Db 735 LeuSerLeuMetGlyLeuIleHisLysProGlnValPheLysValAlaIleAlaGlyAla 754
Qy 576 CCACTCACTCTGGAATCTTATGATACAGGATACAGGAACGCTTATATGGTCCACCT 635
Db 755 ProValThrValTrpMetAlaTyrAspThrGlyTyrThrGluArgTyrMetAspValPro 774
Qy 636 GACAGATGAACAGGCTATTACTTAGATCTGGCCATCTGCGCATGCAAGCAAGAAATCC 695
Db 775 GluAsnGlnHisGlyTyrGluAlaGlySerValAlaLeuHisValGluLysLeuPro 794
Qy 696 TCTGAACCAAAATCGTTTACTGCTCTTACATGGTTTCTGGATGAGATGTCATTTTGA 755
Db 795 AsnGluProAsnArgLeuLeuIleLeuHisGlyPheLeuAspGluAsnValHisPhePhe 814
Qy 756 CATACCAATATATTACTAGTATTTTGTAGTGGCTGGAAAGCCATATGATTTACAGATC 815
Db 815 HisThrAsnPheLeuValSerGlnLeuIleArgAlaGlyLysProTyrGlnLeuGlnIle 834
Qy 816 TATCCTCAGGAGACACAGCATAAGATTCCTGAATCGGAGAACATATATGAAGTCAT 875
Db 835 TyrProAsnGluArgHisSerIleArgCysProGluSerGlyGluHisTyrGluValThr 854
Qy 876 CTTTTCACCTACCTTCAAGAAACCTT 902
Db 855 LeuLeuHisPheLeuGlnGluTyrLeu 863
RESULT 5
Q4SBM6_TETNG
ID Q4SBM6_TETNG PRELIMINARY; PRT; 923 AA.
```

```
AC Q4SBM6; 2005 (T-EMBLrel. 31, Created)
DT 13-SEP-2005 (T-EMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (T-EMBLrel. 31, Last annotation update)
DE Chromosome 15 SCAF14667, whole genome shotgun sequence.
DN (Fragment).
DE ORFNames=GSTENG0020903001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OX NCBI_Taxid=99883;
RN [1]
NUCLEOTIDE SEQUENCE.
RP Jallou O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano B.,
RA Anhouard V., Jubin C., Castellani V., Katinka M., Vacherie B.,
RA Bismont C., Skalli Z., Cattolico L., Poullain J., De Berardinis V.,
RA Ceraud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J.,
RA Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Wolff J.N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Ouetier P., Saurin W., Scarpelli C.,
RA Winkler P., Lander E.S., Weissbach J., Roest Crollius H.,
RT Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RL the early vertebrate proto-karyotype.;
RN [2]
NUCLEOTIDE SEQUENCE.
RP GenomeScope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CAAB01014667; CAG01956.1; -; Genomic_DNA.
FT NON TER 1 1
FT NON TER 923 923
SQ SEQUENCE 923 AA; 105211 MW; 1397023B2004D009 CRC64;
```

```
Alignment Scores:
Pred. No.: 1,13e-91 Length: 923
Score: 1084.50 Matches: 199
Percent Similarity: 66.6% Conservative: 52
Best Local Similarity: 52.8% Mismatches: 49
Query Match: 50.7% Indels: 77
DB: Gaps: 2
```

US-10-825-632-4 (1-1197) x Q4SBM6_TETNG (1-923)

```
Qy 3 TTTGAAGGACCAAGACCTCCCTTTAGAGCATCACCTGACCTAGCTAGTACGTAAAT 62
Db 547 PheGlnGlyThrArgAspThrProLeuGluHisLeuTyrValValSerTyrAspSer 566
Qy 63 CTGGAGAGTGTCACAGGCTGACTGACCGTGGCTACTCACATTTCTGTCATCAGT--- 119
Db 567 ProGlyAspValValArgLeuThrLysProGlyPheSerHisSerCysSerValSerGln 586
Qy 119 ----- 119
Db 587 ValLysLysSerLeuGlnSerAspTyrPhePheAsnTyrSerSerIleThrLeuProLeu 606
Qy 120 -----CAGCAGCTGTGACTCTTTTATAAGTAAGTAAGTAAACACAG 158
Db 607 SerLeuSerSerPheIleTrpGlnAsnPheAspPhePheValSerHisTyrSerVal 626
Qy 159 AAGAATCCACATGTGTGCTCCCTTTTACAAGCTATCAAGCTCTGGAAGATGACCCCACTGC 218
Db 627 CysThrProCysValHisValTyrLysLeuAsnSerSerGluSerAspProLeuHis 646
Qy 219 AAAACAAAGGAATTTTGGGCCACCATTTTGGATTTCAGCAGGTCCTCTCTCTGACTATACT 278
```



```
Db 371 LeuSerLeuMetGlyLeuLeuHisLysProGlnValPheLysValAlaLeuAlaGlyAla 390
Qy 576 CCACTACTCTGTGGATCTTCTATGATACAGGATACAGGAACGTTATATATGGTCAACCT 635
Db 391 ProValThrValTrpMetAlaTyAspThrGlyTyThrGluAUGTyMetAspValPro 410
Qy 636 GACAGAAATGAACAGGGCTATTACTTAGGATCTGGGATGCGAAGCAAGAAAGTTCCCC 695
Db 411 GluAsnAsnGlnHisGlyTyGluAlaGlySerValAlaLeuHisValGluLysLeuPro 430
Qy 696 TCTGAACCAATCGTTTACTGCTTACATGCTTCTTACATGCTTCTGGATGGAATGTCATTTTGA 755
Db 431 AsnGlnProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnValHisPhePhe 450
Qy 756 CATACCAAGTATATTACTGATGTTTTTTAGTACGGGCTGGAAAGCCCATATGATTTA----- 809
Db 451 HisThrAsnPheLeuValSerGlnLeuLeuAlaGlyLysProTyrglnLeuGlnVal 470
Qy 810 -----CAGATCTATCTCTAGGAGAGACACAGCATTAAGAGTTCTCT 848
Db 471 AlaLeuProProValSerProGlnIleTyProAsnGluAUGHisSerIleArgCysPro 490
Qy 849 GAATCGGGAGACATATGATGATGCTTTTGGCACTACCTTCAAGAAACCTT 902
Db 491 GluSerGlyGluHisTyrgluValThrLeuLeuHisPheLeuGlnGluTyLeu 508

RESULT 8
ID Q7PTT8_ANOGA PRELIMINARY; PRT; 621 AA.
AC Q7PTT8;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DE ENSANGP00000015447 (Fragment).
GN ORFNames-ENSANG00000012958;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Culicidae;
OC Anophelinae; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
RT "Anopheles gambiae re-annotation.";
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
CC -! CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAB01008287; EAA03335.3; -; Genomic_DNA.
DR MEROPS; S09.016; -.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001375; Peptidase_S9.
DR InterPro; IPR002469; Peptidase_S9B.
DR Pfam; PF00930; DpPiv N; I.
DR Pfam; PF00326; Peptidase_S9; 1.
FT NON_TER 1
FT NON_TER 621
SQ SEQUENCE 621 AA; 70546 MW; F6EA8463A343BBF3 CRC64;

Alignment Scores:
Pred. No.: 9.2e-59 Length: 621
Score: 730.50 Matches: 154
Percent Similarity: 52.3% Conservative: 42
Best Local Similarity: 41.1% Mismatches: 92
```

```
Query Match: 34.1% Indels: 87
DB: 2 Gaps: 4
US-10-825-632-4 (1-1197) x Q7PTT8_ANOGA (1-621)
Qy 3 TTTCAAGGCAACCAAGACTCCCTTTAGAGCATCACTCTAGCTAGTACGTTACGTTAAAT 62
Db 258 PheMetGlyLeuArgGlnThrProLeuGluLysHisLeuTyValValSerLeuAlaGln 277
Qy 63 CTTGAGAGGTGACCAAGGCTGACTGACCGCTGCTACTCACATTTCTTGTGTCATCAGTCAG 122
Db 278 ProAsnGlnLeuArgLeuLeuThrMetProGlyTySerPheThrValGluPheAsnAsp 297
Qy 123 CACTGTGACTTCTTTATAGTAAATAGTATAGTACCAAGAAATCCACACTGTGTGCTT 182
Db 298 AspCysThrLeuPheLeuGlnThrTyCysAsnIleSerThrLeuProSerTrpGluLeu 317
Qy 193 TACAAGCTA-----TCAAGTCTCGAAGTACGCCCACTTGCACAAACAAAGAAATT 233
Db 318 ValArgIleAlaHisAspSerAsnThrAlaAsnGlyAsnGlyCys----- 332
Qy 234 TGGGCCACCATTTTGGATTTCAGCAGGCTCTCTTCTGACTACTACTCTCT----- 281
Db 333 -----SerHisGlyProThrProProThrProThrProIleAspAlaLeu 346
Qy 281 ----- 281
Db 347 ArgLeuCysSerValGlyTyLeuThrGluGlyGlyProSerGluAsnThrGlnTyAsn 366
Qy 282 CCAGAAATTTCTCTTTTGAAGTACTACTGGATTTCATTTTATGGGATGCTCTACAG 341
Db 367 ProSerIleHisSerProGlnIleSerSerGlyAspValLeuTyAlaMetValPheLys 386
Qy 342 CCTCATGATCTACAGCCTGGAAGAAATATCTACTGTGCTCTCATATATATGCTGCTCT 401
Db 397 ProHisAsnPheMetLeuGlyValLysTyProThrValLeuAsnValTyrglyPro 406
Qy 402 CAG----- 404
Db 407 GluValGlnThrValSerAsnThrPheLysGlyMetArgGlnLeuArgMetHisMetLeu 426
Qy 404 ----- 404
Db 427 AlaSerGlnGlyTyCysValIleCysValAspSerArgGlySerArgHisArgGlyVal 446
Qy 405 -----GGTCAAATAGAAATTCAGCATCAGCATCAGGTGAA 434
Db 447 GluPheGluSerTyrlleArgArgMetGlyThrValGluLeuSerAspGlnValGlu 466
Qy 435 GGACTCCAATATCTAGCTTCTCGATATGATTTTCATTGACTTAGTCTGTGGGCACTCAC 494
Db 467 ValLeuArgIleLeuAlaAspGlnLeuGlyTyrlleAspMetAspArgValAlaIleHis 486
Qy 495 GGCTGTGCTCTATGAGGATACCTCTCCCTGATGGCAATTAATGCAGAGTCAGATATCTTC 554
Db 487 GlyTrpSerTyrglyGlyTyLeuSerLeuMetGlyLeuValGlnTyProGluIlePhe 506
Qy 555 AGGGTGTCTATGCTGGGCCCCAGTCACTCTGTGGATCTCTATGATACAGGATACACG 614
Db 507 LysValSerIleAlaGlyAlaProValThrSerTrpGluTyTyAspThrGlyTyThr 526
Qy 615 GAACGTTATATGGTCCACCTCGACCAAGATGAACAGGGCTATTACTAGGATCTCTGCC 674
Db 527 GluArgTyMetAspLeuProAspSerAsnArgSerGlyTyAlaAlaGlySerValLeu 546
Qy 675 ATGCAAGCAAGAAAGTTCCCTCTCGAACCAATTCGTTTACTCTCTTACATGTTCTCTG 734
Db 547 AsnTyrlleGlnLysPheProAspGluAspAsnArgLeuLeuIleHisGlyLeuLeu 566
Qy 735 GATGAGATGTCCATTTTTCACATACCATATATTTACTGAGTTTCTTGTAGTGGGCTGGA 794
Db 567 AspGluAsnValHisPheHisThrSerGlnLeuValSerArgLeuValArgAlaAsn 586
```


OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Nematocera; Culicidae;
 OC Anophelinae; Anopheles.
 OX NCBI_TaxID=180454;
 RN (1)
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=PEST;
 RG The Anopheles gambiae Sequence Committee;
 RC "Anopheles gambiae re-annotation."
 RT Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 RL (2)
 RN NUCLEOTIDE SEQUENCE.
 RC STRAIN=PEST;
 RG The Anopheles gambiae Sequence Committee;
 RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
 CC -!- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 DR EMBL; AA801008679; EAL40717.1; -; Genomic DNA.
 DR GO; GO:0003824; F: catalytic activity; IEA.
 DR GO; GO:0008236; F: serine-type peptidase activity; IEA.
 DR GO; GO:0006508; P: proteolysis and peptidolysis; IEA.
 DR InterPro; IPR001375; Peptidase_S9.
 DR InterPro; IPR002469; Peptidase_S9B.
 DR InterPro; IPR000379; Ser_estrs.
 DR Pfam; PF00930; DPPIV_N; 1.
 DR Pfam; PF00326; Peptidase_S9; 1.
 FT NON_TER 1
 SQ SEQUENCE 740 AA; 84151 MW; D59F131BBA095CCB CRC64;
 Alignment Scores:
 Pred. No.: 7,13e-57 Length: 740
 Score: 710.50 Matches: 152
 Percent Similarity: 54.4% Conservative: 44
 Best Local Similarity: 42.2% Mismatches: 93
 Query Match: 33.2% Indels: 71
 DB: 2 Gaps: 4
 US-10-825-632-4 (1-1197) x Q5T7K8_ANOGA (1-740)
 QY 3 TTTCAGGCGACCAAGACTCCCTTTAGAGCATCACTGTAGTAGTCACTGTAAT 62
 DB 392 PheMetGluLeuArgGluThrProLeuGluLysHisLeuTyrValValSerLeuAlaGln 411
 QY 63 CCTCGAGGAGTCAGAGCGTCACTGACCGTGGCTACTACATCTCTGCTGATCAGTCAG 122
 DB 412 ProAsnGlnLeuArgLeuLeuThrMetProGlyTyrSer----- 424
 QY 123 CACTGTGACTCTTTTAAAGTAAGTATAGTAAC-----CAGAGAAGATCCACAC 170
 DB 425 -----PheThrValGluPheAsnAspValSerValPheGluArgLysGlyAla 440
 QY 171 TGTGTGTCCTTTACAGCTATCAAGTCTCTGAGATGACCCAACTTGCAAAACAAAGAA 230
 DB 441 CysProAsnThrAlaAsnGlyAsnGlyCysSerHisGlyProThrProProThrPro 460
 QY 231 TTTTGGGCGCCATTGTTGGATTACAGAGGTCTCTTCTCTGACTACTACTCTCT- 281
 DB 461 IleAspAlaLeuArgLeuCysSerValGlyTyrLeuThrGluGlyGlyProSerGluAsn 480
 QY 282 -----CCAGAAATTTTCTCTTTTGAAGTACTACTGGATTTCATATTGATGGG 329
 DB 481 ThrGlnTyrAsnProSerIleHisSerProGlnIleSerSerGlyAspValLeuTyrAla 500
 QY 330 ATGCTCTACAGCTCATGATCTACAGCTCGCTGGAAGAATATCTCTACTGCTGCTCATTA 369
 DB 501 MetValPheLysProHisAsnPheMetMetLeuGlyValLysTyrProThrValLeuAsnVal 520
 QY 390 TATGGTGGTCTCTCAG----- 404
 DB 521 TyrGlyGlyProGluValGlnThrValSerAsnThrPheLysGlyMetArgGlnLeuArg 540
 QY 404 ----- 404

Db 541 MethisMetLeuAlaSerGlnGlyTyrCysValIleCysValAspSerArgGlySerArg 560
 QY 405 -----GGTCAATAGAAATTCAC 422
 DB 561 HisArgGlyValGluPheGluSerTyrIleArgCysArgMetGlyThrValGluLeuSer 580
 QY 423 CATCAGGTGAAGGACTCCCAATATCTAGCTTCTCGATATGATATTCATTCAGTACGT 482
 DB 581 AspGlnValGluValLeuArgIleLeuAlaAspGlnLeuGlyTyrIleAspMetAspArg 600
 QY 483 GTGGGCATCCACGGCTGCTCTATGAGGATCACTCTCCCTGATGGCATTATGAGAGG 542
 DB 601 ValAlaIleHisGlyTyrSerTyrGlyTyrLeuSerLeuMetGlyLeuValGlnTyr 620
 QY 543 TCAGATATCTTCAGGGTGTCTATGCTGGGGCCCACTCTCTGATCTCTTCTATGAT 602
 DB 621 ProGluLeuPheLysValSerIleAlaGlyAlaProValThrSerTyrGluTyrTyrAsp 640
 QY 603 ACAGGATACACGGAACGTTATATGGTCACTCCCTGACCAAGATGAACAGGGCTATTACTTA 662
 DB 641 ThrGlyTyrThrGluArgTyrMetAspLeuProAspSerAsnArgSerGlyTyrAlaAla 660
 QY 663 CGATCTGTGGCCATGCAAGCAGCAAAAGTTCCTCTGACCAAAATCGTTACTGCTCTTA 722
 DB 661 GlySerValLeuAsnTyrIleGlnLysPheProAspGluAspAsnArgLeuLeuIle 680
 QY 723 CATGTTCTCTGAGTGAAGATGTCATTTGCACATACCAGTATATTTACTGAGTTT 782
 DB 681 HisGlyLeuIleAspGluAsnValHisPheHisHisThrSerGlnLeuValSerArgLeu 700
 QY 783 GTGAGGGCTGGAAGCCATATGATTTACAGATCTATCTCTCAGGAGAGACACAGATAAGA 842
 DB 701 ValArgAlaAsnLysProTyrGlnLeuGlnValTyrProAsnGluArgHisSerLeuArg 720
 QY 843 GTTCTCTGAATCGGGAGAACATTTATGAACTGCATCTTTTCCACTACTCTTCAGAAAACCTT 902
 DB 721 AsnLeuGluAlaSerLysHisTyrGluThrLysLeuLeuSerPheLeuGlnAsnHisLeu 740
 RESULT 11
 ID Q9VC20_DROME PRELIMINARY; PRT; 1053 AA.
 AC Q9VC20; Q8IH07;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
 DE CG3744-PA, isoform A (CG3744-pc, isoform c) (LD33755p).
 GN ORFNames=CG3744;
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN (1)
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Chervy J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,
 RA Flossler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,

RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Markulov G., Milshina N.V., Mobaraj C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Spadling A.C., Stapleton M., Skupski M.P., Smith T.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yan R.-F., Zaveri J.S., Zhan N., Zhang G., Zhao Q., Zhao L.,
RA Zheng X.H., Zhong E.W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*.";
RN Science 287:12185-12195 (2000).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426065; PubMed=12537568;
RA Celiniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA George R.A., Hoskins R.A., Laverty T., Muzny D.M., Nelson C.R.,
RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA Svirskas R., Taber P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weinstein G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT "Finishing a whole-genome shotgun: release 3 of the *Drosophila*
RT *melanogaster* euchromatic genome sequence.";
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079 (2002).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426070; PubMed=12537573;
RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J.W., Svirskas R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celiniker S.E.;
RT "The transposable elements of the *Drosophila melanogaster* euchromatin:
RT a genome perspective.";
RL Genome Biol. 3:RESEARCH0084.1-RESEARCH0084.20 (2002).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426069; PubMed=12537572;
RA Miara S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,
RA Smith C.B., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celiniker S.E., de Grey A.D.N.J., Drysdale R.A.,
RA Harris N.L., Richter J., Ruben S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.W., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a
RT systematic review.";
RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22 (2002).
RN [5]
RP NUCLEOTIDE SEQUENCE.
RX Berkeley *Drosophila* Genome Project;
RA Celiniker S., Carlson J., Wan K., Pfeiffer B., Frise E., George R.,
RA Hoskins R., Stapleton M., Pacleb J., Park S., Svirskas R., Smith E.,
RA Yu C., Rubin G.;
RT "*Drosophila melanogaster* release 4 sequence.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [6]
RP NUCLEOTIDE SEQUENCE.
RX FlyBase;
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
RN [7]
RP NUCLEOTIDE SEQUENCE.
RX STRAIN=Berkeley;
RA Stapleton M., Brokatein P., Hong L., Agbayan A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dreanek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,

RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celiniker S.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
CC -|- INTERACTION:
CC Q9VZRO:CG12016; NbExp=1; IntAct=EBI-105926; EBI-152819;
DR EMBL; AE003749; AAF56357.2; -; Genomic_DNA.
DR EMBL; BT001499; AAN71254.1; -; mRNA.
DR IntAct; Q9VC20; -;
DR MEROPS; S09.016; -;
DR Ensembl; CG3744; Drosophila melanogaster.
DR FlyBase; FBgn0039240; CG3744.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0003824; F:diacylglycerol 4-phosphate 4-kinase activity; IEA.
DR GO; GO:0004774; F:diacylglycerol 4-phosphate 4-kinase activity; IEA.
DR GO; GO:0006508; P:protein catabolic process and peptidolysis; IEA.
DR InterPro; IPR001375; Peptidase_S9.
DR InterPro; IPR002469; Peptidase_S9.
DR InterPro; IPR000379; Ser. esterase.
DR Pfam; PF00930; DPEIV_N; I.
DR Pfam; PF00326; Peptidase_S9; 1.
DR SQ SEQUENCE 1053 AA, 118053 MW, C94AA663AB464577 CRC64;
Alignment Scores:
Pred. No.: 2,86e-56 Length: 1053
Score: 704.50 Matches: 150
Percent Similarity: 53.2% Conservative: 44
Best Local Similarity: 41.1% Mismatches: 96
Query Match: 32.9% Indels: 75
DB: 2 Gaps: 6
US-10-825-632-4 (1-1197) x Q9VC20_DROME (1-1053)
QY 3 TTGAAAGGCACCAAGACTCCCTTTAGAGCATACCTGTGACGAGTACGAGTACGTAAT 62
DB 694 PheValGlyLeuArgAspThrProLeuGluLysHisLeuTyrValValSerLeuGluArg 713
QY 63 CCTGGAGAGTGACAAAGGCTGACTGACCGTGGCTACTCACCATTCTTGC----- 110
DB 714 ProGluHisIleArgLeuLeuThrGluProGlyTyrSerTyrLeuValGluPheAspAsp 733
QY 111 -----TGCATCAGTCAGCACTGTGACTCTCTTTATAGTAAGTAT 149
DB 734 HisPheAsnAspLeuPheProIlePheSerGlnGlnCysLysLeuMetLeuValTyr 753
QY 150 AGTAAC---CAGAAAGATCCACACTGTGTGTCCTTTACAAAGCTATCAAGTCTCGAAGAT 206
DB 754 CysAsnIleGlnArgLeuProSerCysLysValMet-----ArgVal 767
QY 207 GACCCAACTTGCAAAACAAAGGAATTTTGGGCCACCATTTTGGATTGAGCAGGT----- 260
DB 768 AsnGlnThrCysSerAsnGlyGlyValAsnGlyIleGlnIleSerLeuValGlyTyrLeu 787
QY 261 -----CCTCTTCTGACTATATCTCTCCAGAAATTTCTCTTTGAAAGT 305
DB 788 HisGluGlyGlyLysProGluProGlnTyrCys---ProGlnIlePheSerProGlnLeu 806
QY 306 ACTACTGGATTACATTGATGGATGCTCTACAGCTCATGATCTACACCTCGGAAG 365
DB 807 ProSerGlyAspIleValTyrAlaMetValPheLysProHisAsnGlyGlyVal 826
QY 366 AATATCTCTACTGCTGCTTCATATATGCTGCTCTCAG----- 404
DB 827 LysTyrProThrValLeuAsnValTyrGlyGlyProGluValGlnThrValAsnAsnThr 846
QY 404 ----- 404
DB 847 PheLysGlyLysHisGlnLeuArgMetHisMetLeuAlaGlnGlyTyrCysValIle 866
QY 404 ----- 404
DB 867 CysIleAspSerArgGlySerArgHisArgGlyLysArgPheGluSerHisIleArgGly 886

Qy 405 -----CGTCAAAATGAAATGACGATCAGGTGGAGGACTCCCAATATCTAGCTTCGGA 458
Db 887 ArgMetGlyGlnValGluThrAspGlnValAspAlaLeuArgSerAspGln 906
Qy 459 TATGATTTCATTCAGTCTAGATCGTGGCCATCCAGCGCTGCTCTATGGAGGATACCTC 518
Db 907 LeuGlyTyrIleAspMetAspArgValAlaIleHisGlyIleTyrPserTyrGlyGlyTyrLeu 926
Qy 519 TCCCTGATGCATTAATACGAGAGTGCAGATATCTTCAGGGTTGCTATTGCTGGGGCCCCA 578
Db 927 SerLeuMetGlyLeuValGlnTyrProLysIlePheLysValAlaIleAlaGlyAlaPro 946
Qy 579 GTCACTCTGTGGATCTTATGATACAGATACAGGAACGCTTATATGCTGGTCCACCTGAC 638
Db 947 ValThrAsnTrpGluTyrAspThrGlyTyrThrGluArgTyrMetAspMetProGln 966
Qy 639 CAGAATCAACAGCGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAAGTCTCCCTCT 698
Db 967 AsnAsnGluAlaGlyTyrSerAlaGlySerValLeuGluTyrValAsnSerPheProGlu 986
Qy 699 GAACCAATCGTTTACTGCTCTTACATGCTTTCTCGATGAGATGTCCTCAATTTTGACAT 758
Db 987 GluAspLysArgLeuLeuLeuIleHisGlyLeuIleAspGluAsnValHisPheCysHis 1006
Qy 759 ACAGTATATATCTAGTCTTTTATGAGGCTGGAAAGCCATATGATTACAGATCTAT 818
Db 1007 ThrSerArgLeuIleSerAlaLeuAsnLysAlaAsnLysProTyrGluValHisLeuPhe 1026
Qy 819 CCTCAGAGAGACACAGCATAGAGTCTCTGATCGGAGAACATATTAGACTGCATCTT 878
Db 1027 ProGluGluArgHisSerLeuArgAsnLeuGluSerAsnLysAsnTyrGluThrLysLeu 1046
Qy 879 TTGCACCTACCTTCAA 893
Db 1047 LeuSerPheLeuGln 1051
RESULT 12
Q9VC19 DROME
ID Q9VC19 DROME PRELIMINARY; PRT: 1113 AA.
AC Q9VC19;
DT 01-MAY-2006 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE CG3744-PB, isoform B.
GN Name=CG3744; ORFNames=CG3744;
OS Drosophila melanogaster (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers J.R., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Frankoch C., Baldwin D.,
RA Ballev R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bertram B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glaeser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,

Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Matti B., McIntosh T.C., McLéod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426065; PubMed=12537568;
RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA George R.A., Hoskins R.A., Lavery T., Muzny D.M., Nelson C.R.,
RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weinstein G., Scher S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT "Finishing a whole-genome shotgun: release 3 of the Drosophila
melanogaster euchromatic genome sequence.";
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426070; PubMed=12537573;
RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J.W., Svirskas R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celniker S.E.;
RT "The transposable elements of the Drosophila melanogaster euchromatin:
a genomics perspective.";
RL Genome Biol. 3:RESEARCH0084.1-RESEARCH0084.20(2002).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Bertram B.P.,
RA Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
systematic review.";
RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
RN [5]
RP NUCLEOTIDE SEQUENCE.
RX Berkeley Drosophila Genome Project;
RA Celniker S., Carlson J., Wan K., Pfeiffer B., Frise E., George R.,
RA Hoskins R., Stapleton M., Pacleb J., Park S., Svirskas R., Smith E.,
RA Yu C., Rubin G.;
RT "Drosophila melanogaster release 4 sequence.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [6]
RP NUCLEOTIDE SEQUENCE.
RG FlyBase;
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB003749; AAF56356.2; -; Genomic_DNA.
DR MEROPS; S09.016; -;
DR Ensembl; CG3744; Drosophila melanogaster.
DR FlyBase; FBgn0039240; CG3744.
DR GO; GO:0003824; F: catalytic activity; IEA.
DR GO; GO:0008236; F: serine-type peptidase activity; IEA.
DR GO; GO:0006508; F: proteolysis and peptidolysis; IEA.
DR InterPro; IPR001375; Peptidase_S9.
DR InterPro; IPR002469; Peptidase_S9B.

```
DR InterPro; IPR000379; Ser_estr.
DR Pfam; PF00930; DPPIV_N; 1.
DR Pfam; PF00326; Peptidase_S9; 1.
SQ SEQUENCE 1113 AA; 123881 MW; 20857E3B212DF2E4 CRC64;

Alignment Scores:
Pred. No.: 2,9e-56 Length: 1113
Score: 704.50 Matches: 150
Percent Similarity: 53.2% Conservativity: 44
Best Local Similarity: 41.1% Mismatches: 96
Query Match: 32.9% Indels: 75
DB: 2 Gaps: 6

US-10-825-632-4 (1-1197) x Q9VC19_DROME (1-1113)
QY 3 TTGGAAGGACCAAGACTCCCTTTAGACGATCACCTGTACGTAGTACGTACGTAAAT 62
DB 754 PheValGlyLeuArgAspThrProLeuGluHisLeuTyrValValSerLeuGluArg 773
QY 63 CCTGGAGAGGTGACAGGCTGACTGACCGCTGGCTACTCACATCTCTGC----- 110
DB 774 ProGluHisIleArgLeuLeuThrGluProGlyTyrSerTyrLeuValGluPheAspAsp 793
QY 111 -----TGCATCATGTCAGCAGCTGTGACTCTTTTATAAGTAAGTAT 149
DB 794 HisPheAsnAspLeuPheProIlePheSerGlnGlnCysLeuMetLeuLeuValTyr 813
QY 150 AGTAAC---CAGAGATCCACACTGTGTGCTCCCTTACAGCTATCAAGTCTCTGAAT 206
DB 814 CysAsnIleGlnArgLeuProSerCysLeuValMet-----ArgVal 827
QY 207 GACCCAACTGCAAAACAAAGGAATTTGGCGCCACCATTTGGATTCAGCAGGT----- 260
DB 828 AsnGlnThrCysSerAsnGlyValAsnGlyIleGlnIleSerLeuValGlyTyrLeu 847
QY 261 -----CCTCTTCTGACTATCTCTCCAGAAATTTCTTTTGAAGT 305
DB 848 HisGluGlyCysProGluProGlnTyrCys---ProGlnIlePheSerProGlnLeu 866
QY 306 ACTACTGGATTTACATTTGATGGATGCTCTACAGCTCATGATCTACAGCTCGGAAG 365
DB 867 ProSerGlyAspIleValTyrAlaMetValPheLeuProHisAsnPheGluLeuGlyVal 886
QY 366 AATATCTCTACTGTGTCTATATATATGTTGCTCTCAG----- 404
DB 887 LysTyrProThrValLeuAsnValTyrGlyGlyProGluValGlnThrValAsnThr 906
QY 404 ----- 404
QY 907 PheLysGlyLysHisGlnLeuArgMetHisMetLeuAlaAlaGlnGlyTyrCysValIle 926
QY 404 ----- 404
QY 927 CysIleAspSerArgGlySerArgHisArgGlyLysArgPheGluSerHisIleArgGly 946
QY 405 -----GTCMAATAGAAATTCACGATCAGGTGGAGGACTCCCAATATCTAGCTTCTCGA 458
DB 947 ArgMetGlyGlnValGluLeuThrAspGlnValAspAlaLeuArgSerLeuSerAspGln 966
QY 459 TATGATTTCTATGATCTGCTGGCATCCGCTGCTCTATGAGGATACCTC 518
DB 967 LeuGlyTyrIleAspMetAspArgValAlaIleHisGlyIlePheSerTyrGlyTyrLeu 986
QY 519 TCCTGATGCGCATTAATGACAGGTCAGATATCTTCAGGTTGCTATTGCTGGGGCCCA 578
DB 987 SerLeuMetGlyLeuValGlnTyrProLysIlePheLysValAlaIleAlaGlyAlaPro 1006
QY 579 GTCACTCTGTGGATCTTCTATGATACAGGATACAGGAACTTATATGGTTCACCTGAC 638
DB 1007 ValThrAsnTrpGlnTyrAspThrGlyTyrThrGluArgTyrMetAspMetProGln 1026
QY 639 CAGAATGAACAGGCTATTACTAGGATCTGTGGCCATGCAAGCAGAAAGATTCCCTCT 698
```

```
Db 1027 AsnAsnGluAlaGlyTyrSerAlaGlySerValLeuGluTyrValAsnSerPheProGlu 1046
QY 699 GAACCAAAATCGTTTACTGCTCTTATCATGTTTCCTGGATGAGAAATGTCATTTGCACAT 758
DB 1047 GluAspLysArgLeuLeuLeuIleHisGlyLeuIleAspGluAsnValHisPheCysHis 1066
QY 759 ACCAGTATATTAAGTATGTTTGTAGTGGGCTGGAAGCCATATGATTTACAGATCTAT 818
DB 1067 ThrSerArgLeuIleSerAlaLeuAsnLysAlaAsnLysProTyrGluValHisLeuPhe 1086
QY 819 CCTCAGGAGAGACACAGCATAAGAGTTCTCTGAATCGGAGAACATTTAGAACTCATCTT 878
DB 1087 ProGluGluArgHisSerLeuArgAsnLeuGluSerAsnLysAsnTyrGluThrLysLeu 1106
QY 879 TTGCACTACCTCAA 893
DB 1107 LeuSerPheLeuGln 1111

RESULT 13
ID Q5TXJ2_ANOGA PRELIMINARY; PRT; 557 AA.
AC Q5TXJ2;
DT 01-FEB-2005 (T-EMBLrel. 29, Created)
DT 01-FEB-2005 (T-EMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (T-EMBLrel. 29, Last annotation update)
DE ENSANGP0000029249 (Fragment).
CN ORFNames=ENSANG0000012958.
OS Anopheles gambiae str. PEST;
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Culicidae;
OC Anophelinae; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
RT "Anopheles gambiae re-annotation.";
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAA01008287; EAL42106.1; -; Genomic DNA.
DR GO; GO:0003824; F: catalytic activity; IEA.
DR GO; GO:0008236; F: serine-type peptidase activity; IEA.
DR GO; GO:0006508; P: proteolysis and peptidolysis; IEA.
DR InterPro; IPR001375; Peptidase_S9.
DR InterPro; IPR002469; Peptidase_S9B.
DR InterPro; IPR000379; Ser_estr.
DR Pfam; PF00930; DPPIV_N; 1.
DR Pfam; PF00326; Peptidase_S9; 1.
FT NON_TER 1 1
FT NON_TER 557 557
SQ SEQUENCE 557 AA; 62685 MW; 08AAA0944ACD0808 CRC64;

Alignment Scores:
Pred. No.: 2,99e-50 Length: 557
Score: 639.00 Matches: 143
Percent Similarity: 53.9% Conservativity: 35
Best Local Similarity: 43.3% Mismatches: 84
Query Match: 29.9% Indels: 68
DB: 2 Gaps: 6

US-10-825-632-4 (1-1197) x Q5TXJ2_ANOGA (1-557)
QY 3 TTGAGGACCAAGACTCCCTTTAGACGATCACCTTTAGACGATCACCTTTAGACGATCAAT 62
DB 239 PheMetGlyLeuArgGlnThrProLeuGluLysHisLeuTyrValValSerLeuAlaGln 258
```


Db 576 Ty-Ser-Lys-Asp-Lys-Lys-Tyr-Pro-Thr-Val-Val-Tyr-Val-Tyr-Gly-Gly-Pro-His-Val-Gln 595
Qy 401 -----
Db 596 Ile-Val-Ile-Arg-Asn-Gln-Tyr-Asn-Tyr-Ile-Lys-Gln-His-Tyr-Thr-Asn-Phe-Gly-Phe-Ile-Gln 615
Qy 401 -----
Db 616 Val-Met-Ile-Asp-Asn-Val-Gly-Ser-Ala-Asn-Arg-Gly-Leu-Glu-Phe-Glu-Ser-His-Ile-Arg 635
Qy 402 -----
Db 636 Glu-Lys-Met-Gly-Gln-Val-Glu-Ile-Gly-Asp-Gln-Val-Glu-Gly-Ile-Asn-Tyr-Leu-Val-Gly 655
Qy 456 CGATATGATTTC-----ATTGACATTAGATCGTGTGGGCATCCACGGCTGCTCTATGGA 509
Db 656 Asn---Asp-Ile-Val-Ser-Ile-Asp-Val-Asn-Arg-Ile-Ala-Ile-Ser-Gly-Tyr-Ser-Tyr-Gly 674
Qy 510 GGATACCTCTCCCTGATGCGATTAAATGACAGAGTCCAGATATCTTCAGGGTTGCTATTGCT 569
Db 675 Gly-Tyr-Asn-Ser-Leu-Met-Ala-Ile-Ser-Gln-Arg-Pro-Asp-Val-Phe-Lys-Ile-Ala-Val-Cys 694
Qy 570 GGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGATACACGGAACCTTATATGGGT 629
Db 695 Gly-Val-Pro-Val-Ser-Asp-Tyr-Arg-Leu-Tyr-Asn-Thr-Gly-Tyr-Thr-Glu-Arg-Tyr-Met-Asn 714
Qy 630 CACCTGACAGATGACAGGCTATTACTTACTTGTAGATCTGTGGCATCTGCAGCAGAGAAAG 689
Db 715 Val-Pro-Gln-Asp-Asn-Ile-Asp-Gly-Tyr-Lys-Leu-Gly-Asp-Thr-Thr-His-Tyr-Ser----- 732
Qy 690 TTCCCTCTGAAACCAATCGTTTACTGTCTTACATGTTTCTGGAATGAGATGATCCAT 749
Db 733 Phe-Pro-Thr-Glu-Glu-Asn-Arg-Leu-Leu-Ile-His-Gly-Leu-Gln-Asp-Glu-Asn-Val-His 752
Qy 750 TTGCGCATACCAAGTATATTACTAGTGTGTTTGTAGAGGCTGGAAAGCCATATGATTTA 809
Db 753 Phe-Ser-Asn-Thr-Ile-Tyr-Ile-Ile-Asp-His-Leu-Thr-Lys-Thr-Gln-Lys-Pro-Tyr-Ile-Leu 772
Qy 810 CAGATCTATCTCCAGAGAGACACAGCATACAGATTCCTGAATCGGAGAACATTTATGAA 869
Db 773 Lys-Thr-Leu-Pro-Asn-Glu-Arg-His-Gly-Val-Arg-Asn-Thr-Asp-Asn-Arg-Ile-Tyr-Ile-Gly 792
Qy 870 CTGCATCTTTTGACCTACCTTCAAGAAACCTT 902
Db 793 Leu-Phe-Val-Ile-Asn-His-Leu-Leu-Lys-Asn-Leu 803

RESULT 15
Q965K3 CABEL
ID Q965K3 CABEL PRELIMINARY; PRT; 927 AA.
AC
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Dipeptidyl peptidase four (Iv) family protein 3, isoform b.
GN Name=dpf-3; ORFNames=K02F2.1;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxId=6239;
[1]
RN
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RG The C. elegans sequencing consortium;
RT "Genome sequence of the nematode C. elegans: a platform for
RL investigating biology."
RL Science 282:2012-2018 (1998).
DR EMBL; AF043699; AAK84627.1; -; Genomic_DNA.
DR MEROPS; S09.016; -.
DR Ensembl; K02F2.1; Caenorhabditis elegans.
DR WormBase; WBGene0001056; dpf-3.
DR WormPep; K02F2.1b; CE28837.
DR GO; 0016020; C:membrane; IEA.

DR GO; 0003824; F: catalytic activity; IEA.
DR GO; 0004274; F: dipeptidyl-peptidase IV activity; IEA.
DR GO; 0006508; F: proteolysis and peptidolysis; IEA.
DR InterPro; IPR001375; Peptidase_S9.
DR InterPro; IPR002469; Peptidase_S9B.
DR InterPro; IPR000379; Ser_estrs.
DR Pfam; PF00930; DPPIV_N; 1.
DR Pfam; PF00326; Peptidase_S9; 1.
KW Complete proteome.
SQ SEQUENCE 927 AA; 105366 MW; 6907C6AAEF829D0F CRC64;

Alignment Scores:
Pred. No.: 1,98e-37 Length: 927
Score: 502.00 Matches: 122
Percent Similarity: 48.4% Conservative: 56
Best Local Similarity: 33.2% Mismatches: 104
Query Match: 23.5% Indels: 86
DB: Gaps: 14

US-10-825-632-4 (1-1197) x Q965K3 CABEL (1-927)
Qy 3 TTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTAGTGTAGTTCAGTTACGTAAT 62
Db 559 Tyr-Val-Ala-Asn-Glu-Ser-His-Pro-Thr-Glu-Tyr-Phe-Ile---Cys-Val-Ser-His-Tyr-Arg 577
Qy 63 CCTGGAGAGTGACAGGCTGACTGACCTGGCTACTCACATCTCTGTGTCATCAGTCAG 122
Db 578 Thr-Gly-Gln-His-Ala-Gln-Ileu-Thr-Glu-Ser-Gly-Ile----- 589
Qy 123 CACTGTGACTCTTTTATAAGTAAGTATAGTATACCAAGAG-----AATCCACAC 170
Db 590 ---Cys-----Phe-Lys-Ser-Glu-Arg-Ala-Ala-Gly-Lys-Leu-Ala-Leu-Asp-Lys-Asp-His 606
Qy 171 TGTGTGTCCCTTTACAGCTATCA-----ACTCTGAAGATGACCCCACTTGTGCAAAACA 224
Db 607 Gly-Phe-Ala-Cys-Tyr-Met-Thr-Ser-Val-Gly-Ser-Pro-Ala-Glu-----Cys-Arg-Phe 623
Qy 225 AAGCAATTT---TGC-----GACCAACATT 245
Db 624 Tyr-Ser-Phe-Arg-Tyr-Lys-Glu-Asn-Glu-Val-Leu-Pro-Ser-Thr-Val-Tyr-Ala-Ala-Ser-Ile 643
Qy 246 TTGGATTTCACAGCTCTCT---CTTCTGTAC-----TATACCTCTCCAGAAATTTCTCT 296
Db 644 Thr-Val-Ser-Gly-His-Pro-Gly-Gln-Pro-Asp-Leu-His-Phe-Asp-Ser-Pro-Glu-Met-Ile-Glu 663
Qy 297 TTTGAAAGT---ACTACTGGATTTTACATTTGATGGGATGCTCTACAAGCCTCATGATCTA 353
Db 664 Phe-Gln-Ser-Lys-Lys-Thr-Gly-Leu-Met-His-Tyr-Ala-Met-Ile-Leu-Arg-Pro-Ser-Asn-Phe 683
Qy 354 CAGCTCGAAAGAAATATCTCTGTGCTGTTCATATATATGGTGGTCTCT----- 401
Db 684 Asp-Pro-Tyr-Lys-Lys-Tyr-Pro-Val-Phe-His-Tyr-Val-Tyr-Gly-Gly-Pro-Gly-Ile-Gln-Ile 703
Qy 401 ----- 401
Db 704 Val-His-Asn-Asp-Ser-Trp-Ile-Gln-Tyr-Ile-Arg-Phe-Cys-Arg-Leu-Gly-Tyr-Val-Val 723
Qy 401 ----- 401
Db 724 Val-Phe-Ile-Asp-Asn-Arg-Gly-Ser-Ala-His-Arg-Gly-Ile-Glu-Phe-Glu-Arg-His-Ile-His 743
Qy 402 -----CAGGTCAAATAGAAATTGACGATCAGCTGAGGAGGAGCTCCATATCTAGCTTCT 455
Db 744 Lys-Lys-Met-Gly-Thr-Val-Glu-Val-Glu-Asp-Gln-Val-Glu-Gly-Leu-Gln-Met-Leu-Ala-Glu 763
Qy 456 CGATAT---GATTTCATTGACTAGATCGTGTGGGCATCCACGGCTGCTCTATGAGGA 512
Db 764 Arg-Thr-Gly-Gly-Phe-Met-Asp-Met-Ser-Arg-Val-Val-His-Gly-Tyr-Ser-Tyr-Gly-Gly 783
Qy 513 TACCTCTCCCTGATGCGATTAAATGACAGAGTTCAGATATCTTCAGGGTTGCTATTGCTGGG 572
Db 784 Tyr-Met-Ala-Leu-Gln-Met-Ile-Ala-Lys-His-Pro-Asn-Ile-Tyr-Arg-Ala-Ile-Ala-Gly 803

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: May 2, 2006, 01:34:19 ; Search time 7.33538 Seconds
(without alignments)
2698.234 Million cell updates/sec

Title: US-10-825-632-4
Perfect score: 2140
Sequence: 1 atttgaaggcaccacaagac.....ttactataaaaaaaaaaaaaa 1197

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 1144120

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlp
-Q/abs/ABSWEB.spool/US10825632/runat_01052006_105950_3311/app_query.fasta_1
-DB=Issued_Patents_AA -QMT=fastan -SUFFIX=n2p.ra -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptco -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-HOST=abs02p -USER=US10825632 @CGN 1.1.131 @runat_01052006_105950_3311
-NCPU=6 -ICPU=3 -NO MAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAEXT=0.5 -FGAPOP=6
-FGAEXT=7 -YGAPOP=10 -YGAEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents AA.*

- 1: /cgn2_6/ptodata/1/iaa/5 COMB.pep.*
- 2: /cgn2_6/ptodata/1/iaa/6 COMB.pep.*
- 3: /cgn2_6/ptodata/1/iaa/H COMB.pep.*
- 4: /cgn2_6/ptodata/1/iaa/PCTUS COMB.pep.*
- 5: /cgn2_6/ptodata/1/iaa/RE COMB.pep.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1680	78.5	310	2	US-09-794-236-4
2	1680	78.5	310	2	US-10-070-464-3
3	1645.5	76.9	882	2	US-09-976-674-1
4	1645.5	76.9	882	2	US-10-070-464-1
5	1377.5	64.4	465	2	US-10-070-464-5
6	1093.5	51.1	863	2	US-09-976-674-3
7	1093.5	51.1	892	2	US-09-976-674-23
8	1093.5	51.1	892	2	US-09-976-674-27
9	1086.5	50.8	360	2	US-10-070-464-7
10	1086.5	50.8	879	2	US-09-976-674-33
11	1086.5	50.8	879	2	US-09-976-674-35
12	867	40.5	690	2	US-09-976-674-7

13	739	34.5	658	2	US-09-976-674-19	Sequence 19, Appl
14	739	34.5	661	2	US-09-976-674-11	Sequence 11, Appl
15	629.5	29.4	832	2	US-09-976-674-29	Sequence 29, Appl
16	629.5	29.4	832	2	US-09-976-674-31	Sequence 31, Appl
17	622.5	29.1	819	2	US-09-976-674-37	Sequence 37, Appl
18	622.5	29.1	819	2	US-09-976-674-39	Sequence 39, Appl
19	471	22.0	613	2	US-09-976-674-21	Sequence 21, Appl
20	397	18.6	981	2	US-09-902-540-16812	Sequence 16812, A
21	393.5	18.4	732	2	US-09-518-550-30	Sequence 30, Appl
22	376	17.6	676	2	US-09-518-550-42	Sequence 42, Appl
23	376	17.6	723	2	US-09-518-550-29	Sequence 29, Appl
24	375	17.5	710	2	US-09-518-550-28	Sequence 28, Appl
25	337.5	15.8	766	2	US-10-002-593-6	Sequence 6, Appl
26	337.5	15.8	766	2	US-09-949-016-6146	Sequence 6146, Ap
27	337.5	15.8	766	2	US-10-423-714-6	Sequence 6, Appl
28	333.5	15.6	755	4	PCT-US93-07923-3	Sequence 3, Appl
29	333.5	15.6	759	4	PCT-US93-07923-2	Sequence 2, Appl
30	333.5	15.6	766	1	US-08-230-491A-3	Sequence 3, Appl
31	333.5	15.6	766	1	US-08-619-280A-3	Sequence 3, Appl
32	333.5	15.6	766	1	US-08-940-391-3	Sequence 3, Appl
33	333.5	15.6	766	2	US-09-794-236-1	Sequence 1, Appl
34	333.5	15.6	766	2	US-09-265-606-3	Sequence 3, Appl
35	333.5	15.6	775	2	US-09-949-016-10450	Sequence 10450, A
36	329.5	15.4	766	2	US-09-518-550-27	Sequence 27, Appl
37	327	15.3	771	2	US-09-462-284-2	Sequence 2, Appl
38	327	15.3	771	2	US-09-079-592-2	Sequence 2, Appl
39	326.5	15.3	737	2	US-09-502-540-11421	Sequence 11421, A
40	317.5	14.8	323	2	US-09-270-767-45296	Sequence 45296, A
41	314	14.7	931	2	US-09-079-592-11	Sequence 11, Appl
42	310	14.5	818	2	US-09-462-845-3	Sequence 3, Appl
43	310	14.5	818	2	US-10-402-312-3	Sequence 3, Appl
44	310	14.5	818	2	US-10-401-437-3	Sequence 3, Appl
45	310	14.5	818	2	US-10-402-067-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1
US-09-794-236-4
; Sequence 4, Application US/09794236
; Patent No. 6337069
; GENERAL INFORMATION:
; APPLICANT: Grouzmann, Eric
; APPLICANT: Lacroix, Jean-Silvain
; APPLICANT: Monod, Michel
; TITLE OF INVENTION: Method of Treating Rhinitis and Sinusitis
; FILE REFERENCE: 81985/276823
; CURRENT APPLICATION NUMBER: US/09/794,236
; CURRENT FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 310
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-794-236-4

Alignment Scores:
Pred. No.: 7.93e-191 Length: 310
Score: 1680.00 Matches: 310
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 78.5% Indels: 0
DB: 2 Gaps: 0

US-10-825-632-4 (1-1197) x US-09-794-236-4 (1-310)

QY	3	TTTGAAGCACAAGACTCCCTTTAGACATCACTGCTAGTACGTACGTTACGTAAT 62
Db	1	PheGluGlyThrIysAspSerProLeuGluHisLeuTyValValSerTyValAsn 20
QY	63	CCTGGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCATCTTGTGCTGATGCTAG 122

Db 21 ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysValSerGln 40
Qy 123 CACTGTGACTTCTTTATAGTAAGTATAGTAACACAGAGAAATCCACACTGTGTGCTCCCTT 182
Db 41 HisCysAspPhePheIleSerLysTyrSerAenGlnLysAenProHisCysValSerLeu 60
Qy 183 TACAAGCTATCAAGTCTCAAGATGACCAACTTGCAGAAACAAAGGAATTTTGGGCAC 242
Db 61 TyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPheTrpAlaThr 80
Qy 243 ATTTTGGATTTCAGCAGTCTCTCTCTGACTATATCTCTCCAGAAATTTTCTTTTGA 302
Db 81 IleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePheSerPheGlu 100
Qy 303 AGTACTACTGGATTACATTGATGGATGCTCTACAGCCTCATGATCTACAGCCTGGA 362
Db 101 SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly 120
Qy 483 GTGGGCATCCACGGCTGGTCTTATAGTAAGTATAGTAACACAGAGAAATTTTGGGCAC 542
Db 161 ValGlyIleHisGlyTrpSerTyrGlyLysLeuSerLeuMetAlaLeuMetGlnArg 180
Qy 543 TCAGATATCTTCAGGGTTCCTTATGCTGGGGCCCCAGTCACCTCTGTGGATCTTCTATGAT 602
Db 181 SerAspIlePheArgValAlaIleAlaGlyAlaProValThrLeuTrpIlePheTyrAsp 200
Qy 603 ACAGGATACACGGAACGTTATATGGTTCACCTGACACAGAAATGAACGGGCTATTACTTA 662
Db 201 ThrGlyTyrThrGluArgTyrMetGlyHisProAspGlnAenGluGlnGlyTyrTyrLeu 220
Qy 663 GGATCTGTGGCCATCAAGCAGAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCTTA 722
Db 221 GlySerValAlaMetGlnAlaGluLysPheProSerGluProAenArgLeuLeuLeu 240
Qy 723 CATGGTTTCTGGATGAGATGTCATTTGACATACAGATCTATCTCTCAGGAGACACAGATAAGA 842
Db 241 HisGlyPheLeuAspGluAenValHisPheAlaHisThrSerIleLeuLeuSerPheLeu 260
Qy 903 GGATCAGTATTCGCTCTTAAAGTGATA 932
Db 301 GlySerArgIleAlaAlaLeuLysValIle 310

RESULT 2

US-10-070-464-3
; Sequence 3, Application US/10070464
; Patent No. 6881564
; GENERAL INFORMATION:
; APPLICANT: ABBOTT, Catherine Anne
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES
; FILE REFERENCE: GH-007
; CURRENT APPLICATION NUMBER: US/10/070,464
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: PCT/AU00/01085
; PRIOR FILING DATE: 2000-09-11
; PRIOR APPLICATION NUMBER: AU PQ5709
; PRIOR FILING DATE: 2000-02-18

; PRIOR APPLICATION NUMBER: AU PQ2762
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 310
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-070-464-3
Alignment Scores: 7.93e-191 Length: 310
Pred. No.: 1680.00 Matches: 310
Score: 100.0% Conservativeness: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 78.5% Indels: 0
DB: 2 Gaps: 0
US-10-825-632-4 (1-1197) x US-10-070-464-3 (1-310)
Qy 3 TTTGAAGGCACCAAGAGCTCCCTTTAGAGCATCACCTGTAGCTAGTCAGTTACGTAAT 62
Db 1 PheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSerTyrValAen 20
Qy 63 CCTGGAGAGGTGACAGGCTGACTGACCGTGGCTACTCACATTCTTCTGCTCAGTCAG 122
Db 21 ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysValSerGln 40
Qy 123 CACTGTGACTTCTTTATAGTAAGTATAGTAACACAGAGAAATTTTGGGCAC 182
Db 41 HisCysAspPhePheIleSerLysTyrSerAenGlnLysAenProHisCysValSerLeu 60
Qy 183 TACAAGCTATCAAGTCTCAAGATGACCAACTTGCAGAAACAAAGGAATTTTGGGCAC 242
Db 61 TyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPheTrpAlaThr 80
Qy 243 ATTTTGGATTTCAGCAGTCTCTCTCTGACTATATCTCTCCAGAAATTTTCTTTGAA 302
Db 81 IleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePheSerPheGlu 100
Qy 303 AGTACTACTGGATTACATTGATGGATGCTCTACAGCCTCATGATCTACAGCCTGGA 362
Db 101 SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly 120
Qy 363 AAGAAATATCTACTGTCTGTCTATATATGGTGGTCTCTCAGGTCACAAATAGAAATGAC 422
Db 121 LysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnGlyGlnIleGluIleAsp 140
Qy 423 GATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTTCATTGACTTAGATCGT 482
Db 141 AspGlnValGluGlyLeuGlnTyrLeuAlaSerArgTyrAspPheIleAspLeuAspArg 160
Qy 483 GTGGGCATCCACGGCTGGTCTTATAGTAAGTATAGTAACACAGAGAAATTTTGGGCAC 542
Db 161 ValGlyIleHisGlyTrpSerTyrGlyLysLeuSerLeuMetAlaLeuMetGlnArg 180
Qy 543 TCAGATATCTTCAGGGTTCCTTATGCTGGGGCCCCAGTCACCTCTGTGGATCTTCTATGAT 602
Db 181 SerAspIlePheArgValAlaIleAlaGlyAlaProValThrLeuTrpIlePheTyrAsp 200
Qy 603 ACAGGATACACGGAACGTTATATGGTTCACCTGACACAGAAATGAACGGGCTATTACTTA 662
Db 201 ThrGlyTyrThrGluArgTyrMetGlyHisProAspGlnAenGluGlnGlyTyrTyrLeu 220
Qy 663 GGATCTGTGGCCATCAAGCAGAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCTTA 722
Db 221 GlySerValAlaMetGlnAlaGluLysPheProSerGluProAenArgLeuLeuLeu 240
Qy 723 CATGGTTTCTGGATGAGATGTCATTTGACATACAGATCTATCTCTCAGGAGACACAGATAAGA 782
Db 241 HisGlyPheLeuAspGluAenValHisPheAlaHisThrSerIleLeuLeuSerPheLeu 260
Qy 783 GTGAGGCTCGAAAGCCATATGATTTACAGATCTATCTCTCAGGAGACACAGATAAGA 842

Db 261 ValArgAlaGlyLysProTyrAspLeuGlnIleTyrProGlnGluArgHisSerIleArg 280
Qy 843 GTTCCTGAATCGGAGAACATTATGAACCTGATCTTTTGCACTACCTTCAAGAAACCTT 902
Db 281 ValProGluSerGlyGluHisTyrGluLeuHisLeuHisTyrLeuGlnGluAsnLeu 300
Qy 903 GGATCAGTATTGCTGCTCTAAAGTGATA 932
Db 301 GlySerArgIleAlaLeuLysValIle 310
RESULT 3
US-09-976-674-1
; Sequence 1, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; PRIOR FILING DATE: 2001-10-12
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 882
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-1
Alignment Scores:
Pred. No.: 1,84e-186 Length: 882
Score: 1645.50 Matches: 310
Percent Similarity: 86.4% Conservative: 0
Best Local Similarity: 86.4% Mismatches: 0
Query Match: 76.9% Indels: 49
DB: 2 Gaps: 1
US-10-825-632-4 (1-1197) x US-09-976-674-1 (1-882)
Qy 3 TTTGAAGCACCAGACCTCCCTTTAGAGCATCACCTGTACGTAGTCAGTTACGTAAAT 62
Db 524 PheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSerTyrValAsn 543
Qy 63 CTTGGAGAGGTGACAGGCTGACTGACCGTGCTACTCACATTTCTGCTGCATCAGTCAG 122
Db 544 ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysIleSerGln 563
Qy 123 CACTGTGACTTTCTTATAGTAACTAGTATAGTACAGAGATCCACACTGTGTCCCTT 182
Db 564 HisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeu 583
Qy 183 TACAAGCTATCAAGTCTCAAGATGACCACTGCAAAACAAAGGAATTTGGGCACC 242
Db 584 TyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPheTpaIaThr 603
Qy 243 ATTTTGGATTACAGCGTCTCTTCTGACTATATCTCTCCAGAAATTTCTCTTTTGAA 302
Db 604 IleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePheSerPheGlu 623
Qy 303 AGTACTACTGGATTACATTGATGGGATGCTTACAGCCTCATGATCTACAGCCTGGA 362
Db 624 SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly 643
Qy 363 AAGAAATATCTTACTGTGCTTATATATATGTTGCTCTCAG----- 404
Db 644 LysLysTyrProThrValLeuPheIleTyrGlyProGlnValGlnLeuValAsnAsn 663
Qy 404 ----- 404

Db 664 ArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrValVal 683
Qy 404 ----- 404
Db 684 ValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLys 703
Qy 405 -----GGTCAAAATAGAAATTCAGATCAGGTGGAGGACTCCAATATATAGCTTCT 455
Db 704 TyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyrLeuAlaSer 723
Qy 456 CGATATGATTTTCACTTACTAGATCGTGTGGGCATCCAGGCTGCTCTATGGAGGATAC 515
Db 724 ArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrPheSerTyrGlyGlyTyr 743
Qy 516 CTCTCCCTGATGCGATTAAATGCAGAGCTCAGATATCTTCAGGGTGTCTATTGCTGGGGCC 575
Db 744 LeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAlaGlyAla 763
Qy 576 CCAGTCACTCTGTGGATCTTCTATATGATACAGGATACAGGAACGTTATATGGTCCACT 635
Db 764 ProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMetGlyHisPro 783
Qy 636 GACCAGAAACAGAGGCTATTACTTAGCATCTGTGGCCATCGACGACGAGAAAGTTCCCC 695
Db 784 AspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGluLysPhePro 803
Qy 696 TCTGAACCAATCGTTTACTGCTCTTACATGCTTCTGGATGAGATGTCATTTTGCAT 755
Db 804 SerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnValHisPheAla 823
Qy 756 CATACCAGTATATTACTAGTATTTTGTAGTGGCGGTGGAAAGCCATATGATTTACAGATC 815
Db 824 HisthrSerIleLeuLeuSerPheLeuValAlaGlyLysProTyrAspLeuGlnIle 843
Qy 816 TATCTCAGAGAGACACAGCATATAGATTTCTGTGATCGGAGACATATGAACTGCAT 875
Db 844 TyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyrGluLeuHis 863
Qy 876 CTTTGTGCACTACCTTCAAGAAACCTTGGATCACGTTATTGCTGCTCTCTAAAAGTGATA 932
Db 864 LeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaLeuLysValIle 882
RESULT 4
US-10-070-464-1
; Sequence 1, Application US/10070464
; Patent No. 6881564
; GENERAL INFORMATION:
; APPLICANT: ABBOTT, Catherine Anne
; APPLICANT: GORRELL, Mark Douglas
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES
; FILE REFERENCE: GH-007
; CURRENT APPLICATION NUMBER: US/10/070,464
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: PCT/AU00/01085
; PRIOR FILING DATE: 2000-09-11
; PRIOR APPLICATION NUMBER: AU PQ5709
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: AU PQ2762
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 882
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-070-464-1
Alignment Scores:
Pred. No.: 1,84e-186 Length: 882
Score: 1645.50 Matches: 310
Percent Similarity: 86.4% Conservative: 0
Best Local Similarity: 86.4% Mismatches: 0

```
Query Match: 76.9% Indels: 49
DB: 2 Gaps: 1
US-10-825-632-4 (1-1197) x US-10-070-464-1 (1-882)

Qy 3 TTTGAAGGACCAAAAGACTCCCTTTAGAGCATCACCTGTACGTPAGTCAGTTACGTAAT 62
Db 524 PheGluGlyThrLysAspSerProLeuGluHisLeuTyValValSerTyValAen 543
Qy 63 CTGGAGAGGTGCAAGGCTGACTGACCGTGGCTACTCACATCTTCTGCTCATCAGTCAG 122
Db 544 ProGlyGluValThrArgLeuThrAspArgGlyTySerHisSerCysCysIleSerGln 563
Qy 123 CACTGTGACTTCTTTATAGTAAGTATAGTAACAGAGAATCCACACTGTGTGCTCCCTT 182
Db 564 HisCysAspPhePheIleSerLysTySerAenGlnLysAenProHisCysValSerLeu 583
Qy 183 TACAAGCTATCAAGTCTGGAAGATGACCAACTTGGCAAAACAAAGGAATTTGGGCCACC 242
Db 584 TyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPheTrpAlaThr 603
Qy 243 ATTTTGGATTACAGGCTCCTCTTCTGACTATCTCTCCAGAAATTTCTCTTTGAA 302
Db 604 IleLeuAspSerAlaGlyProLeuProAspTyThrProGluIlePheSerPheGlu 623
Qy 303 AGTACTACTGGAATTTACATTTGATGGATGCTCTACAGCTCATGATCTACAGCTGGA 362
Db 624 SerThrGlyPheThrLeuTyGlyMetLeuTyLysProHisAspLeuGlnProGly 643
Qy 363 AGAATATCTACTGCTGCTGTCTATATATGCTGCTCTCG----- 404
Db 644 LysLysTyProThrValLeuPheIleTyGlyProGlnValGlnLeuValAen 663
Qy 404 ----- 404
Db 664 ArgPheLysGlyValLysTyPheArgLeuAenThrLeuAlaSerLeuGlyTyValVal 683
Qy 404 ----- 404
Db 684 ValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLys 703
Qy 405 -----GGTCAAATAGAATTGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCT 455
Db 704 TyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyLeuAlaSer 723
Qy 456 CGATATGATTTCATTGACTTAGATCGTGTGGGCATCCACGCTGGTCTATGGAGGATAC 515
Db 724 ArgTyAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyGlyGlyTy 743
Qy 516 CTCTCCCTGATGSCATTAATGCAGAGTCAGATATCTTCAGGTTGCTATTGCTGGGCC 575
Db 744 LeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAlaGlyAla 763
Qy 576 CCAGTCACCTCTGTGGATCTTCTATGATACAGGATACACGGNAGCTTATATGGTCACTT 635
Db 764 ProValThrLeuTrpIlePheTyAspThrGlyTyThrGluArgTyMetGlyHisPro 783
Qy 636 GACCAGATGAACAGGCTATTACTTAGGATCTGTGGCCATGACAGCAGAAAGTTCCCT 695
Db 784 AspGlnAenGlnGlnGlyTyTyLeuGlySerValAlaMetGlnAlaGluLysPhePro 803
Qy 696 TCTGAACCAAATCGTTTACTGCTTTACATGGTTTCTGATGAGAAATGTCATTTGCA 755
Db 804 SerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAenValHisPheAla 823
Qy 756 CATACAGPATATTACTAGTTTTTTAGTGAGGCTGGAAGCCATATGATTTACAGATC 815
Db 824 HisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyAspLeuGlnIle 843
Qy 816 TATCTCTCAGGAGACACAGCATAGAGTTCTCTGAATCGGAGAACATTTATGAACGTGCAT 875
Db 844 TyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyGluLeuHis 863
```

```
Qy 876 CTTTTCGACTACTCTTCAAGAAAAACCTTGGAYCACCTATTGCTGTCTTAAAAAGTGATA 932
Db 864 LeuLeuHisTyLeuGlnGluAenLeuGlySerArgIleAlaAlaLeuLysValIle 882
RESULT 5
US-10-070-464-5
; Sequence 5, Application US/10070464
; Patent No. 6881564
; GENERAL INFORMATION:
; APPLICANT: ABBOTT, Catherine Anne
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES
; FILE REFERENCE: GH-007
; CURRENT APPLICATION NUMBER: US/10/070,464
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: PCT/AU00/01085
; PRIOR FILING DATE: 2000-09-11
; PRIOR APPLICATION NUMBER: AU PQ5709
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: AU PQ2762
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH 465
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-070-464-5

Alignment Scores:
Pred. No.: 1,08e-154 Length: 465
Score: 1377.50 Matches: 259
Percent Similarity: 83.5% Conservative: 0
Best Local Similarity: 83.5% Mismatches: 0
Query Match: 64.4% Indels: 51
DB: 2 Gaps: 1

US-10-825-632-4 (1-1197) x US-10-070-464-5 (1-465)

Qy 3 TTTGAAGGACCAAAAGACTCCCTTTAGAGCATCACCTGTACGTCAGTTACGTAAT 62
Db 207 PheGluGlyThrLysAspSerProLeuGluHisLeuTyValValSerTyValAen 226
Qy 63 CTGGAGAGGTGCAAGGCTGACTGACCGTGGCTACTCACATCTTCTGTCATCAGTCAG 122
Db 227 ProGlyGluValThrArgLeuThrAspArgGlyTySerHisSerCysCysIleSerGln 246
Qy 123 CACTGTGACTTCTTTATAGTAAGTATAGTAACAGAGAATCCACACTGTGTGCTCCCTT 182
Db 247 HisCysAspPhePheIleSerLysTySerAenGlnLysAenProHisCysValSerLeu 266
Qy 183 TACAAGCTATCAAGTCTGGAAGATGACCAACTTGGCAAAACAAAGGAATTTGGGCCACC 242
Db 267 TyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPheTrpAlaThr 286
Qy 243 ATTTTGGATTACAGGCTCCTCTTCTGACTATCTCTCCAGAAATTTCTCTTTGAA 302
Db 287 IleLeuAspSerAlaGlyProLeuProAspTyThrProGluIlePheSerPheGlu 306
Qy 303 AGTACTACTGGAATTTACATTTGATGGATGCTCTACAAGCTCATGATCTACAGCTGGA 362
Db 307 SerThrGlyPheThrLeuTyGlyMetLeuTyLysProHisAspLeuGlnProGly 326
Qy 363 AAGAAATATCTACTGTGCTGTCTATATATGTTGGTCTCTCAGGTCAAAATAGAAATTTGAC 422
Db 327 LysLysTyProThrValLeuPheIleTyGlyGlyProGln----- 340
Qy 423 GATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTTCATTGACTTAGTCGT 482
Db 340 ----- 340
Qy 483 GTGGGCATCCACGGCTGGTCTCTATGGAGGATACCTTCCCTGATGGCATTAAATGCAGAGG 542
```



```
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 23
; LENGTH: 892
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-23

Alignment Scores:
Pred. No.: 1,05e-120 Length: 892
Score: 1093.50 Matches: 200
Percent Similarity: 69.3% Conservative: 42
Best Local Similarity: 57.3% Mismatches: 58
Query Match: 51.1% Indels: 49
DB: 2 Gaps: 1

US-10-825-632-4 (1-1197) x US-09-976-674-23 (1-892)

QY 3 TTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTACGTAGTCAGTTCAGTAAAT 62
Db 544 PheGlnGlyThrLysAspThrProLeuGluHisLeuTyrValValSerTyrGluAla 563
QY 63 CCTGGAGGTCACAGGCTGACTCCGCTGCTACTACATCTTCTGTCATCAGTCAG 122
Db 564 AlaGlyLeuValArgLeuThrProGlyPheSerHisSerCysSerMetSerGln 583
QY 123 CACTGTGACTCTTTTATAGTAAGTATAGTAACAGAGAAATCCACATCTGTGTCCCTT 182
Db 584 AsnPheAspMetPheValSerHisTyrSerValSerThrProCysValHisVal 603
QY 183 TACAAGCTATCAAGTCCTGAAGATGACCCCACTTGCAGAAACAAAGAAATTTGGGCCACC 242
Db 604 TyrLysLeuSerGlyProAspAspProLeuHisLysGlnProArgPheTrpAlaSer 623
QY 243 ATTTGGATTTCAGCAGGTCCTCTCTGACTATATCTCTCCAGAAATTTCTCTTTGAA 302
Db 624 MetMetGluAlaAlaSerCysProProAspTyrValProProGluLeuPheHisPheHis 643
QY 303 AGTACTACTGGATTACATTTGATGGGATGCTTACAGGCTCATGATCTACAGCTCGA 362
Db 644 ThrArgSerAspValArgLeuTyrGlyMetIleTyrLysProHisAlaLeuGlnProGly 663
QY 363 AAGAAATATCCTACTGTGCTGTTATATATATGTTGCTCTCAG----- 404
Db 664 LysLysHisProThrValLeuPheValTyrGlyGlyProGlnValGlnLeuValAsnAsn 683
QY 404 ----- 404
Db 684 SerPheLysGlyLysLysTyrLeuArgLeuAsnThrLeuAlaSerLeuGlyTyrAlaVal 703
QY 404 ----- 404
Db 704 ValValIleAspGlyArgGlySerCysGlnArgGlyLeuArgPheGluGlyAlaLeuLys 723
QY 405 -----GGTCAATAGAAATTCAGATCAGGTGGAAGGACTCCCAATATCTAGCTTCT 455
Db 724 AsnGlnMetGlyGlnValGluIleGluAspGlnValGluGlyLeuGlnPheValAlaGlu 743
QY 456 CGATATGATTTCATTGACTTATAGATCGTGTGGGCATCCACGGCTGTGCTTATGAGGATAC 515
Db 744 LysTyrGlyPheIleAspLeuSerArgValAlaIleHisGlyTrpSerTyrGlyGlyPhe 763
QY 516 CTCCTCCTGATGGCAATTAATGAGAGGTGAGATATCTTCAGGGTGTCTATTGTGGGGCC 575
Db 764 LeuSerLeuMetGlyLeuIleHisLysProGlnValPheLysValAlaIleAlaGlyAla 783
QY 576 CAGTCACTCTGTGATCTTCTATGATACAGGATACAGGAACCTTATATGGGTCACTCT 635
Db 784 ProValThrValTrpMetAlaTyrAspThrGlyTyrThrGluArgTyrMetAspValPro 803

; Sequence 27, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 27
; LENGTH: 892
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-27

Alignment Scores:
Pred. No.: 1,05e-120 Length: 892
Score: 1093.50 Matches: 200
Percent Similarity: 69.3% Conservative: 42
Best Local Similarity: 57.3% Mismatches: 58
Query Match: 51.1% Indels: 49
DB: 2 Gaps: 1

US-10-825-632-4 (1-1197) x US-09-976-674-27 (1-892)

QY 3 TTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTACGTAGTCAGTTCAGTAAAT 62
Db 544 PheGlnGlyThrLysAspThrProLeuGluHisLeuTyrValValSerTyrGluAla 563
QY 63 CCTGGAGGTCACAGGCTGACTCCGCTGCTACTACATCTTCTGTCATCAGTCAG 122
Db 564 AlaGlyLeuValArgLeuThrProGlyPheSerHisSerCysSerMetSerGln 583
QY 123 CACTGTGACTCTTTTATAGTAAGTATAGTAACAGAGAAATCCACATCTGTGTCCCTT 182
Db 584 AsnPheAspMetPheValSerHisTyrSerValSerThrProCysValHisVal 603
QY 183 TACAAGCTATCAAGTCCTGAAGATGACCCCACTTGCAGAAACAAAGAAATTTGGGCCACC 242
Db 604 TyrLysLeuSerGlyProAspAspProLeuHisLysGlnProArgPheTrpAlaSer 623
QY 243 ATTTGGATTTCAGCAGGTCCTCTCTGACTATATCTCTCCAGAAATTTCTCTTTGAA 302
Db 624 MetMetGluAlaAlaSerCysProProAspTyrValProProGluLeuPheHisPheHis 643
QY 303 AGTACTACTGGATTACATTTGATGGGATGCTTACAGGCTCATGATCTACAGCTCGA 362
Db 644 ThrArgSerAspValArgLeuTyrGlyMetIleTyrLysProHisAlaLeuGlnProGly 663
QY 363 AAGAAATATCCTACTGTGCTGTTATATATATGTTGCTCTCAG----- 404
Db 664 LysLysHisProThrValLeuPheValTyrGlyGlyProGlnValGlnLeuValAsnAsn 683
QY 404 ----- 404
Db 684 SerPheLysGlyLysLysTyrLeuArgLeuAsnThrLeuAlaSerLeuGlyTyrAlaVal 703
QY 404 ----- 404
Db 704 ValValIleAspGlyArgGlySerCysGlnArgGlyLeuArgPheGluGlyAlaLeuLys 723
QY 405 -----GGTCAATAGAAATTCAGATCAGGTGGAAGGACTCCCAATATCTAGCTTCT 455
Db 724 AsnGlnMetGlyGlnValGluIleGluAspGlnValGluGlyLeuGlnPheValAlaGlu 743
QY 456 CGATATGATTTCATTGACTTATAGATCGTGTGGGCATCCACGGCTGTGCTTATGAGGATAC 515
Db 744 LysTyrGlyPheIleAspLeuSerArgValAlaIleHisGlyTrpSerTyrGlyGlyPhe 763
QY 516 CTCCTCCTGATGGCAATTAATGAGAGGTGAGATATCTTCAGGGTGTCTATTGTGGGGCC 575
Db 764 LeuSerLeuMetGlyLeuIleHisLysProGlnValPheLysValAlaIleAlaGlyAla 783
QY 576 CAGTCACTCTGTGATCTTCTATGATACAGGATACAGGAACCTTATATGGGTCACTCT 635
Db 784 ProValThrValTrpMetAlaTyrAspThrGlyTyrThrGluArgTyrMetAspValPro 803
```



```
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 33
; LENGTH: 879
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-33

Alignment Scores:
Pred. No.: 7,14e-120 Length: 879
Score: 1086.50 Matches: 199
Percent Similarity: 69.3% Conservative: 42
Best Local Similarity: 57.2% Mismatches: 58
Query Match: 50.8% Indels: 49
DB: 2 Gaps: 1

US-10-825-632-4 (1-1197) x US-09-976-674-33 (1-879)
QY 6 GAAGGCCAACAGACTCCCTTTAGAGCATCACCTGTACGTAGTCAGTTACGTAAATCCT 65
Db 532 LysGlyThrLysAspThrProLeuGluHisLeuTyrValValSerTyrGluAlaAla 551
QY 66 GGAGAGGTGACAAAGCTCAGCTGACCGTGTACTACATCTCTTGGTGCATCAGTCAGCAC 125
Db 552 GlyGluIleValArgLeuThrThrProGlyPheSerHisSerCysSerMetSerGlnAsn 571
QY 126 TGTGACTCTTTATAAGTAAGTATAGTAAACAGAGAATCCACACGTGTGTCCTTTAC 185
Db 572 PheAspMetPheValSerHisTyrSerValSerThrProCysValHisValTyr 591
QY 186 AAGCTATCAAGTCTCTGAAGATGACCAACTTGCACAAACAAAGAAATTTTGGCCACCAT 245
Db 592 LysLeuSerGlyProAspAspAspProLeuHisGlnProArgPheIlePalaSerMet 611
QY 246 TTGGATTACAGAGTCTCTCTGACTATCTCTCCAGAAATTTCTCTTTGAAAGT 305
Db 612 MetGluAlaAlaSerCysProAspIleValProProGluIlePheHisThr 631
QY 306 ACTACTGGATTACATTTGATGGATCTCTACAGCCTCATGATCTACAGCTGGAAG 365
Db 632 ArgSerAspValArgLeuTyrGlyMetIleTyrLysProHisAlaLeuGlnProGlyLys 651
QY 366 AAATATCTACTGTCTGTTCAATATATGGTGGCTCAG----- 404
Db 652 LysHisProThrValLeuPheValTyrGlyProGlnValGlnLeuValAsnSer 671
QY 404 ----- 404
Db 672 PheLysGlyIleLysTyrLeuArgLeuAsnThrLeuAlaSerLeuGlyTyrAlaValVal 691
QY 404 ----- 404
Db 692 ValIleAspGlyArgGlySerCysGlnArgGlyLeuArgPheGluGlyAlaLeuLysAsn 711
QY 405 -----GGTCAATAGAAATTTGACGATCAGTGAAGGACTCAATATATCTAGCTTCTCGA 458
Db 712 GlnMetGlyGlnValGluIleGluAspGlnValGluGlyLeuGlnPheValAlaGluLys 731
QY 459 TATGATTTCAATGACTTGTGTTGGGCATCCAGCGTGTCTCTATGGAGGATACCTC 518
Db 732 TyrGlyPheIleAspLeuSerArgValAlaIleHisGlyTyrPseTyrGlyPheLeu 751
QY 519 TCCCTGATGGCAATTAAGTCAGAGGTGATGATATCTTCAGGTGCTTATGCTGGGCCCA 578
Db 519 TCCCTGATGGCAATTAAGTCAGAGGTGATGATATCTTCAGGTGCTTATGCTGGGCCCA 578
```

```
Db 752 SerLeuMetGlyLeuIleHisLysProGlnValPheLysValAlaIleAlaGlyAlaPro 771
QY 579 GTCACTCTGTGGATCTTCTATGATACAGGATACAGGACCTTATATGGTGCACCCCTGAC 638
Db 772 ValThrValTrpMetAlaTyrAspThrGlyTyrThrGluArgTyrMetAspValProGlu 791
QY 639 CAGAAATGACAGGGCTATTACTTAGTACTGTGGCCATGCAAGCAGAAAGTTCCCTCT 698
Db 792 AsnAsnGlnHisGlyTyrGluAlaGlySerValAlaLeuHisValGluLysLeuProAsn 811
QY 699 GAACCAATCGTTTACTGCTCTTACATGGTTTCTGGATGAGATGTCATTTTGCACAT 758
Db 812 GluProAsnArgLeuLeuIleHisGlyPheLeuAspGluAsnValHisPhePheHis 831
QY 759 ACCAGTATATTACTAGTTTTTACTAGGGCTGGAAAGCCATATGATTTTACAGATCTAT 818
Db 832 ThrAsnPheLeuValSerGlnLeuIleArgAlaGlyLysProTyrGlnLeuGlnIleTyr 851
QY 819 CCTCAGGAGAGACACAGCATAAGATTTCTGAATCGGAGAACATTTATGAATGCATCTT 878
Db 852 ProAsnGluArgHisSerIleArgCysProGluSerGlyGluHisTyrGluValThrLeu 871
QY 879 TTGCACCTACCTTCAAGAAACCTT 902
Db 872 LeuHisPheLeuGlnGluTyrLeu 879

RESULT 11
US-09-976-674-35
; Sequence 35, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 35
; LENGTH: 879
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-35

Alignment Scores:
Pred. No.: 7,14e-120 Length: 879
Score: 1086.50 Matches: 199
Percent Similarity: 69.3% Conservative: 42
Best Local Similarity: 57.2% Mismatches: 58
Query Match: 50.8% Indels: 49
DB: 2 Gaps: 1

US-10-825-632-4 (1-1197) x US-09-976-674-35 (1-879)
QY 6 GAAGGCCAACAGACTCCCTTTAGAGCATCACCTGTACGTAGTCAGTTACGTAAATCCT 65
Db 532 LysGlyThrLysAspThrProLeuGluHisLeuTyrValValSerTyrGluAlaAla 551
QY 66 GGAGAGGTGACAAAGCTCAGCTGACCGTGTACTACATCTCTTGGTGCATCAGTCAGCAC 125
Db 552 GlyGluIleValArgLeuThrThrProGlyPheSerHisSerCysSerMetSerGlnAsn 571
QY 126 TGTGACTCTTTATAAGTAAGTATAGTAAACAGAGAATCCACACGTGTGTCCTTTAC 185
Db 572 PheAspMetPheValSerHisTyrSerValSerThrProCysValHisValTyr 591
QY 186 AAGCTATCAAGTCTCTGAAGATGACCAACTTGCACAAACAAAGAAATTTTGGCCACCAT 245
Db 592 LysLeuSerGlyProAspAspAspProLeuHisGlnProArgPheIlePalaSerMet 611
QY 246 TTGGATTACAGAGTCTCTCTGACTATCTCTCCAGAAATTTCTCTTTGAAAGT 305
Db 612 MetGluAlaAlaSerCysProAspIleValProProGluIlePheHisThr 631
QY 306 ACTACTGGATTACATTTGATGGATCTCTACAGCCTCATGATCTACAGCTGGAAG 365
Db 632 ArgSerAspValArgLeuTyrGlyMetIleTyrLysProHisAlaLeuGlnProGlyLys 651
QY 366 AAATATCTACTGTCTGTTCAATATATGGTGGCTCAG----- 404
Db 652 LysHisProThrValLeuPheValTyrGlyProGlnValGlnLeuValAsnSer 671
QY 404 ----- 404
Db 672 PheLysGlyIleLysTyrLeuArgLeuAsnThrLeuAlaSerLeuGlyTyrAlaValVal 691
QY 404 ----- 404
Db 692 ValIleAspGlyArgGlySerCysGlnArgGlyLeuArgPheGluGlyAlaLeuLysAsn 711
QY 405 -----GGTCAATAGAAATTTGACGATCAGTGAAGGACTCAATATATCTAGCTTCTCGA 458
Db 712 GlnMetGlyGlnValGluIleGluAspGlnValGluGlyLeuGlnPheValAlaGluLys 731
QY 459 TATGATTTCAATGACTTGTGTTGGGCATCCAGCGTGTCTCTATGGAGGATACCTC 518
Db 732 TyrGlyPheIleAspLeuSerArgValAlaIleHisGlyTyrPseTyrGlyPheLeu 751
QY 519 TCCCTGATGGCAATTAAGTCAGAGGTGATGATATCTTCAGGTGCTTATGCTGGGCCCA 578
Db 519 TCCCTGATGGCAATTAAGTCAGAGGTGATGATATCTTCAGGTGCTTATGCTGGGCCCA 578
```

```
Db 592 LysLeuSerGlyProAspAspAspProLeuHisLysGlnProArgPheTrpAlaSerMet 611
QY 246 TTGGATTTCAGAGGCTCTCTCTCTGACTATCTCTCCAGAAATTTCTCTTTGAAAGT 305
Db 612 MetGluAlaIaSerCysProAspTyrValProGluLeuPheHisPheHisThr 631
QY 306 ACTATCGATTACATTTGATGGATGCTCTCAAGCCCTCATGATCTACAGCCCTGGAAAG 365
Db 632 ArgSerAspValArgLeuTyrGlyMetIleTyrLysProHisAlaLeuGlnProGlyLys 651
QY 366 AAATATCTACTGCTGCTTCATATATATGCTGCTCAG----- 404
Db 652 LysHisProThrValLeuPheValTyrGlyGlyProGlnValGlnLeuValAsnSer 671
QY 404 ----- 404
Db 672 PheLysGlyIleLysTyrLeuArgLeuAsnThrLeuAlaSerLeuGlyTyrAlaValVal 691
QY 404 ----- 404
Db 692 ValIleAspGlyArgGlySerCysGlnArgGlyLeuArgPheGluGlyAlaLeuLysAsn 711
QY 405 -----GGTCAATAGAAATTCAGATCAGGTGGAAGGACTCCATATATCTAGCTTCGCA 458
Db 712 GlnMetGlyGlnValGluIleGluAspGlnValGluGlyLeuGlnPheValAlaGluLys 731
QY 459 TATGATTTTCAATGACTTATGCTGCTGGGCATCCAGCGTGTCTCTATGAGGATACCTC 518
Db 732 TyrGlyPheIleAspLeuSerArgValAlaIleHisGlyTyrPheTyrGlyGlyPheLeu 751
QY 519 TCCTGATGGCATTAATGAGAGGTTCAGATATCTTCAGGGTTGCTATTGCTGGGGCCCCA 578
Db 752 SerLeuMetGlyLeuIleHisLysProGlnValPheLysValAlaIleAlaGlyAlaPro 771
QY 579 GTCACTCTGGGATCTCTATGATACAGATACAGGAAGTTATATGGTCAACCTGCAC 638
Db 772 ValThrValTrpMetAlaTyrAspThrGlyTyrThrGluArgTyrMetAspValProGlu 791
QY 639 CAGAAATGACAGGGCTATTACTTAGGATCTGGCCATCCAGCAGAAAGTCCCTCT 698
Db 792 AsnAsnGlnHisGlyTyrGluAlaGlySerValAlaLeuHisValGluLysLeuProAsn 811
QY 699 GAACCAATCGTTTACTGCTCTTACATGTTTCTGATGAGAATGTCATTTTGCACAT 758
Db 812 GluProAsnArgLeuLeuIleLeuHisGlyPheLeuAspGluAsnValHisPheHis 831
QY 759 ACCAGTATATTAATGAGTCTTTTATGAGGCTGGAAAGCCATATGATTTACAGATCTAT 818
Db 832 ThrAsnPheLeuValSerGlnLeuIleArgAlaGlyLysProTyrGlnLeuGlnIleTyr 851
QY 819 CCTCAGGAGACACACATAGAGTTCTGATCGGAGACATTTATGAACTGCATCTT 878
Db 852 ProAsnGluArgHisSerIleArgCysProGluSerGlyGluHisTyrGluValThrLeu 871
QY 879 TTGCATCTACTCTCAAGAAACCTT 902
Db 872 LeuHisPheLeuGlnGluTyrLeu 879
```

RESULT 12

```
US-09-976-674-7
; Sequence 7, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976, 674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
```

```
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 690
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-976-674-7
```

Alignment Scores:

```
Pred. No.: 8,74e-94 Length: 690
Score: 867.00 Matches: 167
Percent Similarity: 75.9% Conservatives: 0
Best Local Similarity: 75.9% Mismatches: 0
Query Match: 40.5% Indels: 53
DB: 2 Gaps: 1
```

US-10-825-632-4 (1-1197) x US-09-976-674-7 (1-690)

```
QY 3 TTTGAAGGCACCAAGACTCCCTTTAGAGACTCACTCTAGTGTAGTTCAGTTACGTAAT 62
Db 524 PheGluGlyThrLysAspSerProLeuGluHisLysLeuTyrValValSerTyrValAsn 543
QY 63 CTGGAGAGGTGACAAAGCTGACTGACCGTGGCTACTCATCTTCTGCTGATCAGTCA 122
Db 544 ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCysIleSerGln 563
QY 123 CACTGTGACTTCTTTATAAGTATAGTAACCAAGAAATCCACACTGTGTGCTCCCTT 182
Db 564 HisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeu 583
QY 183 TACAAGCTATCAAGTCTCTGAAGATGACCCCAACTTGCACAAACAAAGAAATTTTGGCCACC 242
Db 584 TyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPheTrpAlaThr 603
QY 243 ATTTTGGATTACAGAGTCTCTCTCTGACTATCTCTCCAGAAATTTCTCTTTGAA 302
Db 604 IleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePheSerPheGlu 623
QY 303 AGTACTACTGATTTTACATTTGATGAGATGCTCTCAAGCCTCATGATCTACAGCTGGA 362
Db 624 SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly 643
QY 363 AAGAAATATCTACTGTGCTGTTTCATATATGGTGGTCTCAGGGTCAAAATAGAAATTGAC 422
Db 644 LysLysTyrProThrValLeuPheIleTyrGly----- 655
QY 423 GATCAGGTGGAGAGACTCCAATATCTAGCTTCTCGATATGATTTTCACTTACATCGT 482
Db 655 ----- 655
QY 483 GTGGCATCCACGGCTGGTCTCTATGGAGGATACCTCTCTCTGATGCATTATATGCAGAG 542
Db 655 ----- 655
QY 543 TCAGATATCTTCAGGGTTGCTATTTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGAT 602
Db 656 -----Arg-LeuLeuLeuLeuGlyProGlnSerLeuCysGlySerSerMetIle 671
QY 603 ACAGATACACGGAACGTTATATGGTCAACCCCTGACCAAGAAATGACAGCGCTATTACT 660
Db 671 eGlnAspThrArgAsnValIleTrpValThrLeuThrArgMetAsnArgAlaIleThr 690
```

RESULT 13

```
US-09-976-674-19
; Sequence 19, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
```



```

Db      544 PheGlnGlyThrLysAspThrProLeuGluHisHisLeuTyrValValSerTyrGluAla 563
QY      63  CCTGAGAGGTGACAAAGCTGACCTGCGTGTACTCACATTTCTGTGTCAGTCAGTCTCT 122
Db      564 AlaGlyGluIleValArgLeuThrProGlyPheSerHisSerCysSerMetSerGln 583
QY      123 CACTGTGACTTCTTTAAGTAAGTATAGTAACCAAGAAATCCACACTGTGTGTCCTTT 182
Db      584 AsnPheAspMetPheValSerHisTyrSerSerValSerThrProCysValHisVal 603
QY      183 TACAAGCTTATCAAGTCTCAAGATGACCCAACTTGCAAAACAAAGAAATTTGGGCCACC 242
Db      604 TyrLysLeuSerGlyProAspAspAspProLeuHisGlnProArgPheTrpAlaSer 623
QY      243 ATTTGGATTCAGCAGGTCTCTTCTCTGACTATCTCTCCAGAAAATTTCTCTTTTGA 302
Db      624 MetMetGluAlaAlaSerCysProProAspTyrValProProGluIlePheHisPheHis 643
QY      303 AGTACTACTGGATTTACATTTGATGGATGCTCTCAAGCCTCATGATCTACAGCCTGGA 362
Db      644 ThrArgSerAspValArgLeuTyrGlyMetIleTyrLysProHisAlaLeuGlnProGly 663
QY      363 AAGAAATATCTACTGCTGCTCATATATGCTGCTCTCAG----- 404
Db      664 LysLysHisProThrValLeuPheValTyrGlyGlyProGlnValGlnLeuValAsnAsn 683
QY      404 ----- 404
Db      684 SerPheLysGlyIleLysTyrLeuArgLeuAsnThrLeuAlaSerLeuGlyTyrAlaVal 703
QY      404 ----- 404
Db      704 ValValIleAspGlyArgGlySerCysGlnArgGlyLeuArgPheGluGlyAlaLeuLys 723
QY      405 -----GGTCAATAGAAATTGACGATCAGGTGGAAGGACTCCATATATCTAGCTTCT 455
Db      724 AsnGlnMetGlyGlnValGluIleGluAspGlnValGluGlyLeuGlnPheValAlaGlu 743
QY      456 CDAATATGATTTTCATTGACTTAGATCGTGTGGGCATCCAGCGCTGCTTATGGAGGATAC 515
Db      744 LysTyrGlyPheIleAspLeuSerArgValAlaIleHisGlyTrpSerTyrGlyGlyPhe 763
QY      516 CTCTCCCTGATGGCAATTAATGCAGAGTCAGATATCTTCAGGGTTGCTATTGCTGGGGCC 575
Db      764 LeuSerLeuMetGlyLeuIleHisLysProGlnValPheLysAlaGlnProLeuAlaTyr 783
QY      576 CAGTCACTCTG 587
Db      784 ProProArgLeu 787

```

Search completed: May 2, 2006, 01:42:08
 Job time : 53.6769 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: May 2, 2006, 02:43:09 ; Search time 34.2823 Seconds
(without alignments)
2917.786 Million cell updates/sec

Title: US-10-825-632-4

Perfect score: 2140

Sequence: 1 attttgaaggcaccacaaagac.....ttactataaaaaaaaaaaaaa 1197

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 3735138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p model -DEV=xlp
-Q=/abss/ABSSWEB_spool/US10825632/runat_01052006_105957_3449/app_query.fasta_1
-DB=Published Applications_AA_Main -Qfmt=fastan -SUFFIX=n2p.rapbm
-MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1
-MATRIX=blosum62 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct
-THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext
-HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abss02p
-USER=US10825632 -CGCN_1_503 -runat_01052006_105957_3449 -NCPU=6 -ICPU=3
-NO WMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-XGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications_AA_Main.*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1680	78.5	310	3	US-09-993-959-4
2	1880	78.5	310	4	US-10-825-632-3
3	1645.5	76.9	882	3	US-09-976-674-1
4	1645.5	76.9	882	4	US-10-054-776-2
5	1645.5	76.9	882	4	US-10-170-789-38
6	1645.5	76.9	882	4	US-10-311-035-9
7	1645.5	76.9	882	4	US-10-072-012-622
8	1645.5	76.9	882	4	US-10-415-122-6
9	1645.5	76.9	882	4	US-10-825-632-1
10	1645.5	76.9	882	5	US-10-982-512-1
11	1605.5	75.0	883	4	US-10-072-012-621

12	1377.5	64.4	465	4	US-10-825-632-5	Sequence 5, Appli
13	1372	64.1	580	4	US-10-275-505-2	Sequence 2, Appli
14	1372	64.1	580	6	US-11-140-224-2	Sequence 2, Appli
15	1094.5	51.1	869	4	US-10-415-122-4	Sequence 4, Appli
16	1093.5	51.1	497	3	US-09-833-245-1594	Sequence 1594, Ap
17	1093.5	51.1	689	4	US-10-072-012-620	Sequence 620, App
18	1093.5	51.1	830	4	US-10-415-122-7	Sequence 7, Appli
19	1093.5	51.1	863	3	US-09-976-674-3	Sequence 3, Appli
20	1093.5	51.1	863	4	US-10-072-012-619	Sequence 619, App
21	1093.5	51.1	863	5	US-10-982-512-3	Sequence 3, Appli
22	1093.5	51.1	892	3	US-09-976-674-23	Sequence 23, Appli
23	1093.5	51.1	892	3	US-09-976-674-27	Sequence 27, Appli
24	1093.5	51.1	892	5	US-10-982-512-23	Sequence 23, Appli
25	1093.5	51.1	892	5	US-10-982-512-27	Sequence 27, Appli
26	1093.5	51.1	892	5	US-10-433-757-12	Sequence 12, Appli
27	1093.5	51.1	969	4	US-10-415-122-2	Sequence 2, Appli
28	1089.5	50.9	497	3	US-09-833-245-1593	Sequence 1593, Ap
29	1086.5	50.8	360	4	US-10-825-632-7	Sequence 7, Appli
30	1086.5	50.8	879	3	US-09-976-674-33	Sequence 33, Appli
31	1086.5	50.8	879	3	US-09-976-674-35	Sequence 35, Appli
32	1086.5	50.8	879	5	US-10-982-512-33	Sequence 33, Appli
33	1086.5	50.8	879	5	US-10-982-512-35	Sequence 35, Appli
34	1085.5	50.7	863	4	US-10-072-012-224	Sequence 224, App
35	1085.5	50.7	863	4	US-10-072-012-226	Sequence 226, App
36	1032.5	48.2	508	4	US-10-237-271-3	Sequence 3, Appli
37	1032.5	48.2	508	4	US-10-072-012-623	Sequence 623, App
38	938.5	43.9	312	4	US-10-094-749-1686	Sequence 1686, Ap
39	924	43.2	193	4	US-10-264-237-2115	Sequence 2115, Ap
40	867	40.5	690	3	US-09-976-674-7	Sequence 7, Appli
41	867	40.5	690	5	US-10-982-512-7	Sequence 7, Appli
42	739	34.5	658	3	US-09-976-674-19	Sequence 19, Appli
43	739	34.5	658	5	US-10-982-512-19	Sequence 19, Appli
44	739	34.5	661	3	US-09-976-674-11	Sequence 11, Appli
45	739	34.5	661	5	US-10-982-512-11	Sequence 11, Appli

ALIGNMENTS

RESULT 1

US-09-993-959-4
; Sequence 4, Application US/09993959
; Publication No. US20030165489A1
; GENERAL INFORMATION:
; APPLICANT: Grouzmann, Eric
; APPLICANT: Lacroix, Jean-Silvain
; APPLICANT: Monod, Michel
; TITLE OF INVENTION: Method of Treating Rhinitis and Sinusitis
; FILE REFERENCE: 81985/276823
; CURRENT APPLICATION NUMBER: US/09/993,959
; CURRENT FILING DATE: 2001-11-27
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 310
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-993-959-4

Alignment Scores:
Pred. No.: 4.85e-169 Length: 310
Score: 1680.00 Matches: 310
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 78.5% Indels: 0
DB: 3 Gaps: 0

US-10-825-632-4 (1-1197) x US-09-993-959-4 (1-310)

QY 3 TTGGAAGCACAAGACTCCCTTTAGACATCACCTGTACGTAGTTCAGTTACGTAAAT 62
|||||
Db 1 PheGluGlyThrLysAspSerProLeuGluHisIleuTyValValSerTyValAsn 20
|||||
QY 63 CCTGAGAGGTGACAAAGGCTGACTGACCGTGGCTACTCACATTTCTTGTGTCATGTCAG 122

```
Db 21 ProGlyGluValThrArgLeuThrAspArgGlyThrSerHisSerCysHisSerGln 40
Qy 123 CACTGTGACTCTTTTATAAGTATAGTAACACAGAGAAATCCACACTGTGTGCCCTT 182
Db 41 HisCysAspPhePheIleSerLysThrSerAsnGlnLysAsnProHisCysValSerLeu 60
Qy 183 TACAAGCTATCAAGTCTCAAGATGACCAACTTGTCAAAACAAAGGAATTTTGGGCACC 242
Db 61 TyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPheTAlaThr 80
Qy 243 ATTTTGGATTTCAGCAGCTCTCTCTGACTATATCTCTCCAGAAATTTTCTTTTGA 302
Db 81 IleLeuAspSerAlaGlyProLeuProAspThrProProGluIlePheSerPheGlu 100
Qy 303 AGTACTACTGGATTATACATTTGATGGATGCTCTACAGGCTCATGATCTACAGCTGGA 362
Db 101 SerThrThrGlyPheThrLeuThrGlyMetLeuThrLysProHisAspLeuGlnProGly 120
Qy 363 AAGAAATATCCCTACTGTGCTGTCATATATGTTGCTCTCAGGCTCAAAATAGAAATGAC 422
Db 121 LysLysThrProThrValLeuPheIleThrGlyGlyProGlnGlyGlnIleGluIleAsp 140
Qy 423 GATCAGGTGGAGGACTCCAAATATCTAGCTTCTCGATATGATTTTATGATGCT 482
Db 141 AspGlnValGluGlyLeuGlnThrLysLeuAlaSerArgThrAspPheIleAspLeuAspArg 160
Qy 483 GTGGGATCCACGGCTGTCTTATGAGGATACCTCTCCCTGATGGGCTCAATAGAAATGAC 542
Db 161 ValGlyIleHisGlyTrpSerThrGlyThrLeuSerLeuMetAlaLeuMetGlnArg 180
Qy 543 TCAGATATCTTCAGGTTCTCTATGCTGGGCCCCAGTCACTCTGTGATCTTCTATGAT 602
Db 181 SerAspIlePheArgValAlaIleAlaGlyAlaProValThrLeuTrpIlePheThrAsp 200
Qy 603 ACAGGATACACGGAAGCTTTATATGGGTACCCCTGACCAAGATGAACAGGCTATTACTTA 662
Db 201 ThrGlyThrThrGluArgThrMetGlyHisProAspGlnAsnGluGlnGlyThrLeu 220
Qy 663 GGATCTGTGGCCATGCAAGCAGAAAGTTCCCTCTGAACCAAAATCGTTTACTGCTTTA 782
Db 241 HisGlyPheLeuAspGluAsnValHisPheAlaHisThrSerIleLeuLeuSerPheLeu 260
Qy 783 GTGAGGCTGGAAGCCATATGATTTACAGATCTATCTCTCAGGAGACACAGCATAAGA 842
Db 261 ValArgAlaGlyLysProThrAspLeuGlnIleThrProGlnGluArgHisSerIleArg 280
Qy 843 GTTCTCGAATCGGAGAACATTTATGACTGATCTTTTGGCACTACCTTCAAGAAACCTT 902
Db 281 ValProGluSerGlyGluHisThrGluLeuHisLeuLeuHisThrLeuGlnGluAsnLeu 300
Qy 903 GGATCAGTATTCCTCTCTTAAAGTGATA 932
Db 301 GlySerArgIleAlaAlaLeuLysValIle 310
```

```
RESULT 2
US-10-825-632-3
; Sequence 3, Application US/10825632
; Publication No. US20040191826A1
; GENERAL INFORMATION:
; APPLICANT: ABBOTT, Catherine Anne
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES
; FILE REFERENCE: FCSB-100-Div. 1
; CURRENT APPLICATION NUMBER: US/10/825,632
; CURRENT FILING DATE: 2004-04-15
; PRIOR APPLICATION NUMBER: US 10/070,464
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: PCT/AU00/01085
```

```
; PRIOR FILING DATE: 2000-09-11
; PRIOR APPLICATION NUMBER: AU PQ5709
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: AU PQ2762
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 310
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-825-632-3
Alignment Scores:
Pred. No.: 4,85e-169 Length: 310
Score: 1680.00 Matches: 310
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 78.5% Indels: 0
DB: Gaps: 0
US-10-825-632-4 (1-1197) x US-10-825-632-3 (1-310)
Qy 3 TTTGAGAGGCACCAAGACTCCCTTTAGAGCATCACCTGTAGTACGTAGTACGTAAAT 62
Db 1 PheGluGlyThrLysAspSerProLeuGluHisLeuThrValValSerThrValAsn 20
Qy 63 CTTGGAGAGGTGACAGGCTGACTGACCGTGGCTACTCACATTTCTTCTGCATCAGTCAG 122
Db 21 ProGlyValThrArgLeuThrAspArgGlyThrSerHisSerCysValSerGln 40
Qy 123 CACTGTGACTCTTTTATAAGTATAGTATACACAGAGAAATCCACACTGTGTGCCCTT 182
Db 41 HisCysAspPhePheIleSerLysThrSerAsnGlnLysAsnProHisCysValSerLeu 60
Qy 183 TACAAGCTATCAAGTCTCTGAGAGTACCCCACTTCCCAAAACAAAGGAATTTTGGCCACC 242
Db 61 TyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPheTrpAlaThr 80
Qy 243 ATTTTGGATTTCAGCAGCTCTCTCTGACTATATCTCTCCAGAAATTTTCTTTTGA 302
Db 81 IleLeuAspSerAlaGlyProLeuProAspThrProProGluIlePheSerPheGlu 100
Qy 303 AGTACTACTGGATTATACATTTGATGGATGCTCTCAAGCCTCATGATCTTACAGCCTGGA 362
Db 101 SerThrThrGlyPheThrLeuThrGlyMetLeuThrLysProHisAspLeuGlnProGly 120
Qy 363 AAGAAATATCTACTGTGCTGTCATATATGTTGCTCTCAGGCTCAAAATAGAAATGAC 422
Db 121 LysLysThrProThrValLeuPheIleThrGlyGlyProGlnGlyGlnIleGluIleAsp 140
Qy 423 GATCAGGTGGAGGACTCCAAATATCTAGCTTCTCGATATGATTTTCAATGACTTAGATCGT 482
Db 141 AspGlnValGluGlyLeuGlnThrLysLeuAlaSerArgThrAspPheIleAspLeuAspArg 160
Qy 483 GTGGGATCCACGGCTGTCTTATGAGGATACCTCTCCCTGATGGGCTCAATAGAAATGAC 542
Db 161 ValGlyIleHisGlyTrpSerThrGlyThrLeuSerLeuMetAlaLeuMetGlnArg 180
Qy 543 TCAGATATCTTCAGGTTCTCTATGCTGGGCCCCAGTCACTCTGTGATCTTCTATGAT 602
Db 181 SerAspIlePheArgValAlaIleAlaGlyAlaProValThrLeuTrpIlePheThrAsp 200
Qy 603 ACAGGATACACGGAAGCTTTATATGGGTACCCCTGACCAAGATGAACAGGCTATTACTTA 662
Db 201 ThrGlyThrThrGluArgThrMetGlyHisProAspGlnAsnGluGlnGlyThrLeu 220
Qy 663 GGATCTGTGGCCATGCAAGCAGAAAGTTCCCTCTGAACCAAAATCGTTTACTGCTTTA 722
Db 221 GlySerValAlaMetGlnAlaGluLysPheProSerGluProAsnArgLeuLeuLeuLeu 240
Qy 723 CATGGTTTCTCGATGAGAAATGTCCATTTTGGCACATACAGTATATTAATGAGTTTATA 782
```



```
Db 241 HisGlyPheLeuAspGluAsnValHisPheAlaHisThrSerIleLeuLeuSerPheLeu 260
QY 783 GTGAGGCGTGGAAGCCATATGATTTACAGATCTATCTCCTCAGGAGACACAGCATAAGA 842
Db 261 ValArgAlaGlyLysProTyrAspLeuGlnIleTyrProGlnGluArgHisSerIleArg 280
QY 843 GTTCCTGAATCGGAGAACATTATGAACTGCATCTTTTGCACCTACCTTCAAGAAAACCTT 902
Db 281 ValProGluSerGlyGluHisTyrGluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeu 300
QY 903 GATACAGTATTCTGCTCTAAAGTGATA 932
Db 301 GlySerArgIleAlaAlaLeuLysValIle 310

RESULT 3
US-09-976-674-1
; Sequence 1, Application US/09976674
; Patent No. US20020115843A1
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 882
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-1.

Alignment Scores:
Pred. No.: 3,21e-165 Length: 882
Score: 1645.50 Matches: 310
Percent Similarity: 86.4% Conservatives: 0
Best Local Similarity: 86.4% Mismatches: 0
Query Match: 76.9% Indels: 49
DB: 3 Gaps: 1

US-10-825-632-4 (1-1197) x US-09-976-674-1 (1-882)
QY 3 TTGAAGCCACCAAGACTCCCTTTAGAGCATCTCCTGACGTAGTCAGTTACGTAAAT 62
Db 524 PheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSerTyrValAsn 543
QY 63 CTTGAGAGGTGACAAGCGTGACGTGACCGGTGCTACATCTCTGCTGCATCAGTCAG 122
Db 544 ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCysIleSerGln 563
QY 123 CACTGTGACTCTTTATAAGTATAGTAACCAAGAAATCCACACTGTGTGCTCCTT 182
Db 564 HisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeu 583
QY 183 TACAAGCTATCAAGTCTCTGAAGATGACCCCAACTTGCAAAACAAAGAAATTTGGGCCACC 242
Db 584 TyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPheTrpAlaThr 603
QY 243 ATTTTGGATTACAGAGTCTCTCTCTGACTATCTCTCCAGAAATTTCTCTTTGAA 302
Db 604 IleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePheSerPheGlu 623
QY 303 AGTACTACTGATTATCATTTGATGGATGCTCTCAAGCCTCATGATCTACAGCTGGA 362
Db 624 SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly 643
QY 363 AAGAAATATCTACTGTGCTTTCATATATGTTGGTCTTCAG----- 404
|||||
```

```
Db 644 LysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnValAsnAsn 663
QY 404 ----- 404
Db 664 ArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrValVal 683
QY 404 ----- 404
Db 684 ValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLys 703
QY 405 -----GGTCAAAATAGAAATTTGACGATCAGGTGGAGGACTCCAATATCTAGCTTCT 455
Db 704 TyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyrLeuAlaSer 723
QY 456 CGATATGATTTTCATTGACTTTAGATCGTGTGGGCATCCACGGCTGCTCTATGAGGATAC 515
Db 724 ArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrSerTyrGlyGlyTyr 743
QY 516 CTCTCCCTGATGCGCATTATGACAGGTGACAGATCATCTTCAGGGTTGCTATTGCTGGGCC 575
Db 744 LeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAlaGlyAla 763
QY 576 CCAGTCACTCTGTGGATCTCTATGATACAGGATACACGACGTTATATGGTCACCCCT 635
Db 764 ProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMetGlyHisPro 783
QY 636 GACCAGAAATGAACAGGGCTATTACTTAGGATCTGTGGCCATCAACAGCAGAAAAGTTCCCC 695
Db 784 AspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGluLysPhePro 803
QY 696 TCTGAACCAAAATCGTTTACTGCTCTTACATGGTTTCTGGATGAGAAATGTCATTTTGA 755
Db 804 SerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnValHisPheAla 823
QY 756 CATACCACTATATCTAGTGTGTTTCTAGTGGCGCTGGAAGCCATATGATTTACAGATC 815
Db 824 HisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAspLeuGlnIle 843
QY 816 TATCTCTCAGGAGACACACAGATAGAGTTCTCTGAATCGGAGAACATTATGAACATGCAT 875
Db 844 TyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyrGluLeuHis 863
QY 876 CTTTGTGACCTACCTTCAAGAAAACCTTGGATCAGCTATTGCTGCTCTAAAAGTGATA 932
Db 864 LeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeuLysValIle 882

RESULT 4
US-10-054-776-2
; Sequence 2, Application US/10054776
; Publication No. US20030165818A1
; GENERAL INFORMATION:
; APPLICANT: Mark Robert Edbrooke
; APPLICANT: Alan Peter Lewis
; TITLE OF INVENTION: NOVEL PROTEIN
; FILE REFERENCE: OGI042US
; CURRENT APPLICATION NUMBER: US/10/054,776
; CURRENT FILING DATE: 2002-01-23
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 882
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-054-776-2

Alignment Scores:
Pred. No.: 3,21e-165 Length: 882
Score: 1645.50 Matches: 310
Percent Similarity: 86.4% Conservatives: 0
Best Local Similarity: 86.4% Mismatches: 0
Query Match: 76.9% Indels: 49
DB: 4 Gaps: 1
```

```
US-10-825-632-4 (1-1197) x US-10-054-776-2 (1-882)
Qy 3 TTTGAAGGACCAAAAGACTCCCTTTAGAGCATCACCTGACGTAGTCAGTTCAGTAAAT 62
Db 524 PheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSerTyrValAsn 543
Qy 63 CTGGAGAGGTGACAAAGGTGACTGACCGTGGCTACTCACATTTCTGCTCATCAGTCAG 122
Db 544 ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCysIleSerGln 563
Qy 123 CACTGTGACTCTTTTATAGTAAAGTATAGTAACAGAGAAATCCACACTGTGTGTCCTTT 182
Db 564 HisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeu 583
Qy 183 TACAACCTCAAGTCCTGAAGTACACCACTTGCAGAAATTTTTCGGCCACC 242
Db 584 TyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPheTrpAlaThr 603
Qy 243 ATTTTCGATTGACAGGTCCTCTCTGACTATCTCTCCAGAAATTTTCTCTTTGAA 302
Db 604 IleLeuAspSerAlaGlyProLeuProAspTyrThrProGluLullePheSerPheGlu 623
Qy 303 AGTACTACTGGATTACATTGATGGGATGCTCTACAGCTCATGATCTACAGCTGGA 362
Db 624 SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly 643
Qy 363 AAGAAATATCTACTGTGCTTCATATATATGTTGCTCTCAG----- 404
Db 644 LysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeuValAsnAsn 663
Qy 404 ----- 404
Db 664 ArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrValVal 683
Qy 404 ----- 404
Db 684 ValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLys 703
Qy 405 -----GGTCAAAATAGAAATTCACATCAGGTGGAGGACTCCAATATCTAGTTCT 455
Db 704 TyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyrLeuAlaSer 723
Qy 456 CGATATGATTTCATTCATTAGATCGTGTGGCATCCACGGCTCGTCTATGAGAGATAC 515
Db 724 ArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrSerTyrGlyTyr 743
Qy 516 CTCTCCCTGATGCATTAATGACAGAGTCAGATATCTTCAGGGTTGCTATTGCTGGGCC 575
Db 744 LeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAlaGlyAla 763
Qy 576 CCAGTCACCTGTGGATCTCTATGATACAGATACACGGAACCTTATATGSGTCCACCT 635
Db 764 ProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMetGlyHisPro 783
Qy 636 GACCAGATGAACAGGGCTATTACTTAGATCTGTGGCCATGCAAGCAGAAAGTTCCCC 695
Db 784 AspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGluLysPhePro 803
Qy 696 TCTGAACCAAAATCGTTTACTGCTCTTACATGGTTTCTCGATGAGAAATGTCATTTGCA 755
Db 804 SerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnValHisPheAla 823
Qy 756 CATACCAAGTATATTACTAGTGTTTTGTAGTGGGCTGGAAGCCCATATGATTTACAGATC 815
Db 824 HisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAspLeuGlnIle 843
Qy 816 TATCTCAGGAGACACAGCATAGAGTTCTCGAATCGGAGAACATTTATGAACATGCAT 875
Db 844 TyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyrGluLeuHis 863
Qy 876 CTTTTCAGCTACTCTCAAGAAACCTTGGATCACCTATTGCTGCTCTAAAAGTGATA 932
Db 864 LeuLeuHisTyrLeuGlnGlnLeuLeuGlySerArgIleAlaAlaLeuLysValIle 882
```

```
RESULT 5
US-10-170-789-38
; Sequence 38, Application US/10170789
; Publication No. US20030180930A1
; GENERAL INFORMATION:
; APPLICANT: Rachel E. Meyers
; APPLICANT: Olandt, Peter J.
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Curtis, Rory A. J.
; APPLICANT: Williamson, Mark
; APPLICANT: Weich, Nadine
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASE, PHOSPHATASE,
; TITLE OF INVENTION: AND PROTEASE FAMILY MEMBERS AND USES THEREOF
; FILE REFERENCE: 10448-191001
; CURRENT APPLICATION NUMBER: US/10170,789
; CURRENT FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: US 09/797,039
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: PCT/US01/06525
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/186,061
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 09/882,166
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: PCT/US01/19269
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/212,078
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 09/934,406
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: PCT/US01/26052
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/226,740
; PRIOR FILING DATE: 2000-08-21
; PRIOR APPLICATION NUMBER: US 09/861,801
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: PCT/US01/16549
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: US 60/205,508
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: US 09/801,267
; PRIOR FILING DATE: 2001-03-06
; PRIOR APPLICATION NUMBER: PCT/US01/07138
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: US 60/187,454
; PRIOR FILING DATE: 2000-03-07
; PRIOR APPLICATION NUMBER: US 09/829,671
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: PCT/US01/40483
; PRIOR FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: US 60/197,508
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: US 09/961,721
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: PCT/US01/29904
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: US 60/235,023
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US 10/045,367
; PRIOR FILING DATE: 2001-11-07
; PRIOR APPLICATION NUMBER: US 60/246,561
; PRIOR FILING DATE: 2000-11-07
; PRIOR APPLICATION NUMBER: US 09/801,275
; PRIOR FILING DATE: 2001-03-06
; PRIOR APPLICATION NUMBER: PCT/US01/07074
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: US 60/187,420
; PRIOR FILING DATE: 2000-03-07
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 38
; LENGTH: 882
```

TYPE: PRT
ORGANISM: Homo sapiens
US-10-170-789-38

Alignment Scores:

Pred. No.: 3, 21e-165 Length: 882
Score: 1645.50 Matches: 310
Percent Similarity: 86.4% Conservative: 0
Best Local Similarity: 86.4% Mismatches: 0
Query Match: 76.9% Indels: 49
DB: 4 Gaps: 1

US-10-825-632-4 (1-1197) x US-10-170-789-38 (1-882)

QY 3 TTGAAGCACAAGACTCCCTTTAGACATCACTCTAGCTAGTACGTTACGTAAT 62
DB 524 PheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSerTyrValAsn 543
QY 63 CTGGAGAGTGACAAGGCTGACTGACCGTGGCTACTACATCTTGTGCTGATCAGTCAG 122
DB 544 ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCysLeuSerGln 563
QY 123 CACTGTGACTCTTTATAAGTAAGTATAGTAACCAAGAAATCCACACTGTGTGCTCCTT 182
DB 564 HisCysAspPhePheLeuSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeu 583
QY 183 TACAAGCTATCAAGTCTCTGAAGATGACCCCAACTTGCAAAACAAAGAAATTTTGGGCCACC 242
DB 584 TyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPheTrpAlaThr 603
QY 243 ATTTGGATTGACAGAGTCTCTCTCTGACTATCTCTCCAGAAATTTTCTCTTTGAA 302
DB 604 IleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIlePheSerPheGlu 623
QY 303 AGTACTACTGATTTACATTTGATGATGATGCTCTCAAGCCTCATGATCTACAGCTGGA 362
DB 624 SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly 643
QY 363 AAGAAATCTCTACTGTGCTTTCATATATGTTGGTCTCTCAG----- 404
DB 644 LysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeuValAsnAsn 663
QY 404 ----- 404
DB 664 ArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrValVal 683
QY 404 ----- 404
DB 684 ValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLys 703
QY 405 -----GGTCAATAGAAATTCAGATCAGGTGGAAGGACTCCAATATCTAGCTTCT 455
DB 704 TyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyrLeuAlaSer 723
QY 456 CQATATGATTTCAATGACTTAGATCGTGTGGGCATCCAGCGTGTGCTCTATGAGGATAC 515
DB 724 ArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyrGlyGlyTyr 743
QY 516 CTCTCCCTGATGCAATTAATGCAGAGTCAAGATCTCTCAGGGTTGCTATTGCTGGGGCC 575
DB 744 LeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAlaGlyAla 763
QY 576 CQAGTCACTCTGTGATCTTCTATGATACAGGATACACGGAACGTTATATGGGTCAACCT 635
DB 764 ProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMetGlyHisPro 783
QY 636 GACCAGAAATGAACGGGCTATTACTTAGGATCTGTGGCCATCCAGCAGAAAGTTCCTCC 695
DB 784 AspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGluLysPhePro 803
QY 696 TCTGACCAAAATCGTTTACTGCTCTTACATGTTTCTGATGAGAAATGTCATTTGCA 755
DB 804 SerGluProAsnArgLeuLeuLeuLeuHisGlyPheLeuAspGluAsnValHisPheAla 823

RESULT 6

US-10-311-035-9
Sequence 9, Application US/10311035
Publication No. US20040023243A1

GENERAL INFORMATION:

APPLICANT: INCYTE GENOMICS, INC.
APPLICANT: YUE, Henry
APPLICANT: ELLIOTT, Vicki
APPLICANT: GANDHI, Ameena R.
APPLICANT: LAL, Preeti
APPLICANT: AU-YOUNG, Janice
APPLICANT: TRIBOULEY, Catherine M.
APPLICANT: DELEGANE, Angelo M.
APPLICANT: BAUGHN, Mariah R.
APPLICANT: NGUYEN, Darniel B.
APPLICANT: LEE, Ernestine A.
APPLICANT: HAFALTA, April
APPLICANT: KHAN, Farrah A.
APPLICANT: CHAWLA, Narinder K.
APPLICANT: YAO, Monique G.
APPLICANT: LU, Dying Aina M.
APPLICANT: ARVIZU, Chandra S.
APPLICANT: TANG, Y. Tom
APPLICANT: WALSH, Roderick T.
APPLICANT: AZIMZAI, Yalda
APPLICANT: LU, Yan
APPLICANT: RAMKUMAR, Jayalaxmi
APPLICANT: XU, Yuming
APPLICANT: REDDY, Roopa
APPLICANT: DAS, Depopriya
APPLICANT: KEARNEY, Liam
APPLICANT: KALLICK, Deborah A.
TITLE OF INVENTION: Proteases
FILE REFERENCE: PI-0123 PCT
CURRENT APPLICATION NUMBER: US/10/311,035
CURRENT FILING DATE: 2002-12-10
PRIOR APPLICATION NUMBER: 60/212,336; 60/213,955; 60/215,396; 60/216,821; 60/218,946
PRIOR FILING DATE: 2000-06-16; 2000-06-22; 2000-06-29; 2000-07-07; 2000-07-14
NUMBER OF SEQ ID NOS: 42
SOFTWARE: PERL Program
SEQ ID NO 9
LENGTH: 882
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. US20040023243A1 7160544CD1

US-10-311-035-9

Alignment Scores:

Pred. No.: 3, 21e-165 Length: 882
Score: 1645.50 Matches: 310
Percent Similarity: 86.4% Conservative: 0
Best Local Similarity: 86.4% Mismatches: 0
Query Match: 76.9% Indels: 49
DB: 4 Gaps: 1

US-10-825-632-4 (1-1197) x US-10-311-035-9 (1-882)

QY 3 TTGAAGCACAAGACTCCCTTTAGACATCACTCTAGCTAGTACGTTACGTAAT 62

Sequence 622, Application US/10072012

Publication No. US20040033493A1

GENERAL INFORMATION:

APPLICANT: Tchernev, Velizar

APPLICANT: Spytek, Kimberly

APPLICANT: Zerkhusen, Bryan

APPLICANT: Patturajan, Meera

APPLICANT: Shimkets, Richard

APPLICANT: Li, Li

APPLICANT: Gangolli, Esha

APPLICANT: Padigaru, Muralidhara

APPLICANT: Anderson, David W.

APPLICANT: Rastelli, Luca

APPLICANT: Miller, Charles E.

APPLICANT: Gerlach, Valerie

APPLICANT: Taupier Jr, Raymond J.

APPLICANT: Gusev, Vladimir Y.

APPLICANT: Colman, Steven D.

APPLICANT: Wolenc, Adam R.

APPLICANT: Pena, Carol E. A

APPLICANT: Furtak, Katarzyna

APPLICANT: Grosse, William M.

APPLICANT: Alsebrook II, John P.

APPLICANT: Lepley, Denise K.

APPLICANT: Rieger, Daniel K.

APPLICANT: Burgess, Catherine E.

TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same

FILE REFERENCE: 21402-258

CURRENT APPLICATION NUMBER: US/10/072,012

CURRENT FILING DATE: 2002-01-31

PRIOR APPLICATION NUMBER: 60/265,102

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: 60/265,514

PRIOR FILING DATE: 2001-01-31

PRIOR APPLICATION NUMBER: 60/265,517

PRIOR FILING DATE: 2001-01-31

PRIOR APPLICATION NUMBER: 60/265,412

PRIOR FILING DATE: 2001-01-31

PRIOR APPLICATION NUMBER: 60/265,395

PRIOR FILING DATE: 2001-01-31

PRIOR APPLICATION NUMBER: 60/266,406

PRIOR FILING DATE: 2001-02-02

PRIOR APPLICATION NUMBER: 60/266,767

PRIOR FILING DATE: 2001-02-05

PRIOR APPLICATION NUMBER: 60/267,057

PRIOR FILING DATE: 2001-02-07

PRIOR APPLICATION NUMBER: 60/266,975

PRIOR FILING DATE: 2001-02-07

PRIOR APPLICATION NUMBER: 60/267,459

PRIOR FILING DATE: 2001-02-08

Remaining prior Application data removed - See File Wrapper or PALM.

NUMBER OF SEQ ID NOS: 1391

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 622

LENGTH: 882

TYPE: PRT

ORGANISM: Homo sapiens

US-10-072-012-622

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

3 21e-165

1645 50

86.4%

86.4%

76.9%

4

Length:

Matches:

Conservative:

Mismatches:

Indels:

Gaps:

882

310

0

49

1

US-10-825-632-4 (1-1197) x US-10-072-012-622 (1-882)

Qy

3

TTTGAAGGACCAAAAGACTCCCTTTAGAGCATCACCTGTAGTACGTACGTAAT

62

Db

524

PheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSerTyrValAsn

543

RESULT 7

US-10-072-012-622

```
QY 63 CTGGAGAGGTGACAGGCTGACTGACCGTGGCTACTCACAATCTTCTGCTGCATCAGTCAG 122
Db 544 ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysValSerGln 563
QY 123 CACTGTGACTCTTTTATAGTAAGTATAGTAACACAGAGAATCCACACTGTGTGTCCTCT 182
Db 564 HisCysAspPhePheLeuSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeu 583
QY 183 TACAAGCTATCAAGTCTCAAGATGACCCCACTTGCACAAACAAAGAAATTTTGGGCCACC 242
Db 584 TyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPheTyrAlaThr 603
QY 243 ATTTTGGATTACAGAGCTCTCTCTGACTATATCTCTCCAGAAATTTTCTCTTTGAA 302
Db 604 IleLeuAspSerAlaGlyProLeuProAspTyrThrProGluLysPheSerPheGlu 623
QY 303 AGTACTACTGGATTATACATTTGATGGGATGCTCTACAGGCTCATGATCTACAGCTGGA 362
Db 624 SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly 643
QY 363 AAGAAATATCTACTGCTGCTGTTCATATATATGCTGCTCTCAG----- 404
Db 644 LysLysTyrProThrValLeuPheLeuTyrGlyProGlnValGlnLeuValAsnAsn 663
QY 404 ----- 404
Db 664 ArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrValVal 683
QY 404 ----- 404
Db 684 ValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLys 703
QY 405 -----GGTCAAAATAGAAATTCAGATTCAGGTGGAGAGGCTCCAATATCTAGCTTCT 455
Db 704 TyrLysMetGlyGlnIleGluLeuAspAspGlnValGluGlyLeuGlnTyrLeuAlaSer 723
QY 456 CQATATGATTTCACTTATGATTCGTGGCATCCACGGCTCGTCTCATGAGGATAC 515
Db 724 ArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrSerTyrGlyTyr 743
QY 516 CTCTCCCTGATGATTAAGTATTAAGTATTAAGTATTAAGTATTAAGTATTAAGTATTAAG 575
Db 744 LeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAlaGlyAla 763
QY 576 CQAGTCACTCTGTGGATCTTCTATGATATACAGGATACAGGAACTGATGGGTCCACCT 635
Db 764 ProValThrLeuThrPheThrAspThrGlyTyrThrGluArgTyrMetGlyHisPro 783
QY 636 GACCAGAAATGAACAGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAGTTCCCC 695
Db 784 AspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGluLysPhePro 803
QY 696 TGTGAACCAAAATCGTTTACTGCTCTTATCATGTTCTTCTGGATGAGATGTCATTTTGC 755
Db 804 SerGluProAsnArgLeuLeuLeuLeuHisGlyPheLeuAspGluAsnValHisPheAla 823
QY 756 CATACAGTATATTACTGAGTGTGAGTGTGAGGCTGGAAGCCATATGATTTTACAGATC 815
Db 824 HisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAspLeuGlnIle 843
QY 816 TATCTCTCAGGAGACACAGCATAAGAGTTCTCGAATCGGAGAACATTAATGAAGTCAT 875
Db 844 TyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyrGluLeuHis 863
QY 876 CTTTTCGACTTACCTTCAAGAAACCTTGGATCAGCATATGCTGCTCTTAAAGTGATA 932
Db 864 LeuLeuHisTyrLeuGlnGlnAsnLeuGlySerArgIleAlaAlaLeuLysValIle 882
```

RESULT 8
US-10-415-122-6
; Sequence 6, Application US/10415122
; Publication No. US20040053369A1
; GENERAL INFORMATION:

```
; APPLICANT: THE UNIVERSITY OF SYDNEY
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASHS
; FILE REFERENCE: FP15217
; CURRENT APPLICATION NUMBER: US/10/415,122
; CURRENT FILING DATE: 2003-08-07
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 882
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-415-122-6

Alignment Scores:
Pred. No.: 3,21e-165 Length: 882
Score: 1645.50 Matches: 310
Percent Similarity: 86.4% Conservative: 0
Best Local Similarity: 86.4% Mismatches: 0
Query Match: 76.9% Indels: 49
DB: 4 Gaps: 1

US-10-825-632-4 (1-1197) x US-10-415-122-6 (1-882)
QY 3 TTTGAAGGACCAAGACTCCCTTTAGAGCATCACCTGTAGTACGTACGTACGTAAT 62
Db 524 PheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSerTyrValAsn 543
QY 63 CTGGAGAGGTGACAGGCTGACTGACCGTGGCTACTCACAATCTTCTGTCATCAGTCAG 122
Db 544 ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysValSerGln 563
QY 123 CACTGTGACTCTTTTATAGTAAGTATTAAGTATTAAGTATTAAGTATTAAGTATTAAG 182
Db 564 HisCysAspPhePheLeuSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeu 583
QY 183 TACAAGCTATCAAGTCTCTGAGATGACCACTTGCACAAACAAAGAAATTTTGGGCCACC 242
Db 584 TyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPheTyrAlaThr 603
QY 243 ATTTTGGATTACAGAGCTCTCTCTGACTATATCTCTCCAGAAATTTTCTCTTTGAA 302
Db 604 IleLeuAspSerAlaGlyProLeuProAspTyrThrProGluLysPheSerPheGlu 623
QY 303 AGTACTACTGGATTATACATTTGATGGGATGCTCTACAGGCTCATGATCTACAGCTGGA 362
Db 624 SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly 643
QY 363 AAGAAATATCTACTGCTGCTGTTCATATATGCTGCTCTCAG----- 404
Db 644 LysLysTyrProThrValLeuPheLeuTyrGlyProGlnValGlnLeuValAsnAsn 663
QY 404 ----- 404
Db 664 ArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrValVal 683
QY 404 ----- 404
Db 684 ValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLys 703
QY 405 -----GGTCAAAATAGAAATTCAGATTCAGGTGGAGAGGCTCCAATATCTAGCTTCT 455
Db 704 TyrLysMetGlyGlnIleGluLeuAspAspGlnValGluGlyLeuGlnTyrLeuAlaSer 723
QY 456 CGATATGATTTCACTTATGATTCGTGGCATCCACGGCTCGTCTCATGAGGATAC 515
Db 724 ArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrSerTyrGlyTyr 743
QY 516 CTCTCCCTGATGATTAAGTATTAAGTATTAAGTATTAAGTATTAAGTATTAAGTATTAAG 575
Db 744 LeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAlaGlyAla 763
QY 576 CQAGTCACTCTGTGGATCTTCTATGATATACAGGATACAGGAACTGATGGGTCCACCT 635
Db 764 ProValThrLeuThrPheThrAspThrGlyTyrThrGluArgTyrMetGlyHisPro 783
```

Db 764 PrbValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMetGlyHisPro 783
QY 636 GACCAGATGACAGGGCTATTACTTAGGATCTGTGGCCATGACAGCAAGAAAGTTCCCC 695
Db 784 AspGlnAsnGluGlnGlyTyrLeuGlySerValAlaMetGlnAlaGluLysPhePro 803
QY 696 TCTGAACCAAAATCGTTTACTGCTTACATCGTTTCTGATGAGAAATGTCATTTTGA 755
Db 804 SerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnValHisPheAla 823
QY 756 CATACAGPATATTACTAGTTTATAGTGGGCTGGAAGCCATATGATTTACAGATC 815
Db 824 HisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAspLeuGlnIle 843
QY 816 TATCTCAGGAGACACAGCATAGAGTTCTCGAATCGGGAGAACATTATGAACATGCAT 875
Db 844 TyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyrGluLeuHis 863
QY 876 CTTTTCACATACCTTCAAGAAACCTTGGATCAGTATTGCTGCTCTAAAAGTGATA 932
Db 864 LeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeuLysValIle 882

RESULT 9

US-10-825-632-1
; Sequence 1, Application US/10825632
; Publication No. US20040191826A1
; GENERAL INFORMATION:
; APPLICANT: ABBOTT, Catherine Anne
; APPLICANT: CORRELL, Mark Douglas
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES
; FILE REFERENCE: FCSB-100-Div. 1
; CURRENT APPLICATION NUMBER: US/10/825.632
; CURRENT FILING DATE: 2004-04-15
; PRIOR APPLICATION NUMBER: US 10/070,464
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: PCT/AU00/01085
; PRIOR FILING DATE: 2000-09-11
; PRIOR APPLICATION NUMBER: AU PQ5709
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: AU PQ2762
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 882
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-825-632-1

Alignment Scores
Pred. No.: 3.21e-165 Length: 882
Score: 1645.50 Matches: 310
Percent Similarity: 86.4% Conservative: 0
Best Local Similarity: 86.4% Mismatches: 0
Query Match: 76.9% Indels: 49
DB: 4 Gaps: 1

US-10-825-632-4 (1-1197) x US-10-825-632-1 (1-882)

QY 3 TTTGAAGGCACCAAGACTCCCTTTAGAGATCACCTGTACGTACGTACGTAAAT 62
Db 524 PheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSerTyrValAsn 543
QY 63 CTGAGAGGTGACAAAGCTGACTGACCGTGGCTACTACATCTTTGCTGCATCAGTCAG 122
Db 544 ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysLysSerGln 563
QY 123 CACTGTGACTCTTTTATAGTAAGTATAGTAACCAAGAAATCCACACTGTGTCCCTT 182
Db 564 HisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeu 583
QY 183 TACAAGCTATCAAGTCTGAGATGACCAACTTGAACAACAAAGAAATTTGGGGCCAC 242

Db 584 TyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPheTrpAlaThr 603
QY 243 ATTTTCGATTCAGAGGCTCTCTCTGACTATACTCTCCAGAAATTTCTCTTTTGAA 302
Db 604 IleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePheSerPheGlu 623
QY 303 AGTACTACTGGATTTACATGTTATGGATGCTCTACAAGCCTCATGATCTACAGCCTGGA 362
Db 624 SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly 643
QY 363 AAGAAATATCTACTGTGCTGTATATATATATGTTGCTCTCAG----- 404
Db 644 LysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeuValAsnAsn 663
QY 404 ----- 404
Db 664 ArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrValVal 683
QY 404 ----- 404
Db 684 ValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLys 703
QY 405 -----GGTCAATAGAAATTCACGATCAGGTGGAAGGACTCCAATATCTAGCTTCT 455
Db 704 TyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyrLeuAlaSer 723
QY 456 CGATATGATTCTGACTTACATGCTGTGGGCATCCACGGCTGCTCTATGAGGATAC 515
Db 724 ArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyrGlyLys 743
QY 516 CTCTCCCTGATGCGATTTATGCGAGGTTCAGATATCTTCAGGGTTGCTATTGCTGGGGCC 575
Db 744 LeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAlaGlyAla 763
QY 576 CCAGTCACTCTGTGATCTTCTATGATACAGGATACACGGAACTTATATGGGTCACCCCT 635
Db 764 ProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMetGlyHisPro 783
QY 636 GACCAGATGAACAGGGCTATTACTTAGATCTGTGGCCATGCAAGCAGAAAAAGTTCCCC 695
Db 784 AspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGluLysPhePro 803
QY 696 TCTGAACCAAAATCGTTTACTGCTTACATGTTTCTCTGATGAGAAATGTCATTTTGA 755
Db 804 SerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnValHisPheAla 823
QY 756 CATACAGTATATTACTGAGTTTTTTAGTAGGGCTGGAAGCCATATGATTTACAGATC 815
Db 824 HisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAspLeuGlnIle 843
QY 816 TATCTCAGGAGACACAGCATAGAGTTCTCGAATCGGGAGAACATTATGAACTGCAT 875
Db 844 TyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyrGluLeuHis 863
QY 876 CTTTTCACATACCTTCAAGAAACCTTGGATCAGTATTGCTGCTCTAAAAGTGATA 932
Db 864 LeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeuLysValIle 882

RESULT 10

US-10-982-512-1
; Sequence 1, Application US/10982512
; Publication No. US20050059081A1
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/10/982.512
; CURRENT FILING DATE: 2004-11-05
; PRIOR APPLICATION NUMBER: US/09/976,674
; PRIOR FILING DATE: 2001-10-12

; PRIOR APPLICATION NUMBER: US 60/240,117

; PRIOR FILING DATE: 2000-10-12

; NUMBER OF SEQ ID NOS: 61

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 1

; LENGTH: 882

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-982-512-1

Alignment Scores:

Pred. No.: 3, 21e-165 Length: 882
Score: 1645.50 Matches: 310
Percent Similarity: 86.4% Conservative: 0
Best Local Similarity: 86.4% Mismatches: 0
Query Match: 76.9% Indels: 49
DB: 5 Gaps: 1

US-10-825-632-4 (1-1197) x US-10-982-512-1 (1-882)

Qy 3 TTTGAAGCCACCAAGACTCCCTTTAGAGCATCACTGTAGTCAGTTCAGTTCAGTAAAT 62
Db 524 PheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSerTyrValAsn 543
Qy 63 CTGGAGAGGTGACAGGCTGACTGACCGTGGCTACTACATTCACATTCCTGCTGATCAGTCAG 122
Db 544 ProGluGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCysIleSerGln 563
Qy 123 CACTGTGACTCTTTATAGTATAGTATACACAGAGAGATCCACACTGTGTCTCCCTT 182
Db 564 HisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeu 583
Qy 183 TACAAGCTATCAAGTCTGAGATGACCCCACTTCACAAACAAAGGAATTTTGGCCACC 242
Db 584 TyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPheTrpAlaThr 603
Qy 243 ATTTTGGATTACAGAGTCTCTCTGACTATCTCTCAGAGAAATTTCTCTTTTGA 302
Db 604 IleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIlePheSerPheGlu 623
Qy 303 AdTACTACTGGATTTACATTTGATGGGATGCTCTACAGCCTCATGATCTACAGCCTGGA 362
Db 624 SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly 643
Qy 363 AAGAAATATCTACTGTGCTGTTTCATATATGTTGCTCTCAG----- 404
Db 644 LysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeuValAsnAsn 663
Qy 404 ----- 404
Db 664 ArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrValVal 683
Qy 404 ----- 404
Db 684 ValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLys 703
Qy 405 -----GTCATAATAGAAATTCAGATCAGGTGGAGGACTCCAATATCTAGCTTCT 455
Db 704 TyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyrLeuAlaSer 723
Qy 456 CdATATGATTTCACTTGAATTCGTGTGGGCTCCACGGCTGCTCTCTATGGAGGATAC 515
Db 724 ArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrSerTyrGlyGlyTyr 743
Qy 516 CTTCTCCCTGATGCAATTAATGACAGGTGAGATATCTTCAGGGTTGCTATTGCTGGGGCC 575
Db 744 LeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAlaGlyAla 763
Qy 576 CCAGTCACTCTGGGATCTCTATATGATACAGGATACAGGAACGTTATATGGTCAACCT 635
Db 764 ProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMetGlyHisPro 783
Qy 636 GACCAGATGACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAGTTCCTCCC 695

RESULT 11

US-10-072-012-621

; Sequence 621, Application US/10072012

; Publication No. US20040033493A1

; GENERAL INFORMATION:

; APPLICANT: Tchernev, Velizar

; APPLICANT: Spytek, Kimberly

; APPLICANT: Zerhusen, Bryan

; APPLICANT: Patturajan, Meera

; APPLICANT: Shimkets, Richard

; APPLICANT: Li, Li

; APPLICANT: Gangolli, Baha

; APPLICANT: Padigaru, Muralidhara

; APPLICANT: Anderson, David W.

; APPLICANT: Rastelli, Luca

; APPLICANT: Miller, Charles E.

; APPLICANT: Gerlach, Valerie

; APPLICANT: Taupier Jr, Raymond J.

; APPLICANT: Gusev, Vladimir Y.

; APPLICANT: Colman, Steven D.

; APPLICANT: Wolenc, Adam R.

; APPLICANT: Pena, Carol E. A

; APPLICANT: Furtak, Katarzyna

; APPLICANT: Grosse, William M.

; APPLICANT: Alsobrook II, John P.

; APPLICANT: Lepley, Denise M.

; APPLICANT: Rieger, Daniel K.

; APPLICANT: Burgess, Catherine E.

; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same

; FILE REFERENCE: 21402-258

; CURRENT APPLICATION NUMBER: US/10/072,012

; CURRENT FILING DATE: 2002-01-31

; PRIOR APPLICATION NUMBER: 60/265,102

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: 60/265,514

; PRIOR FILING DATE: 2001-01-31

; PRIOR APPLICATION NUMBER: 60/265,517

; PRIOR FILING DATE: 2001-01-31

; PRIOR APPLICATION NUMBER: 60/265,412

; PRIOR FILING DATE: 2001-01-31

; PRIOR APPLICATION NUMBER: 60/265,395

; PRIOR FILING DATE: 2001-01-31

; PRIOR APPLICATION NUMBER: 60/266,406

; PRIOR FILING DATE: 2001-02-02

; PRIOR APPLICATION NUMBER: 60/266,767

; PRIOR FILING DATE: 2001-02-05

; PRIOR APPLICATION NUMBER: 60/267,057

; PRIOR FILING DATE: 2001-02-07

; PRIOR APPLICATION NUMBER: 60/266,975

; PRIOR FILING DATE: 2001-02-07

; PRIOR APPLICATION NUMBER: 60/267,459

; PRIOR FILING DATE: 2001-02-08

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 1391

```
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 621
; LENGTH: 883
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-072-012-621

Alignment Scores:
  Pred. No.:      5,72e-161      Length:      883
  Score:          1605.50        Matches:      302
  Percent Similarity: 85.0%      Conservative: 3
  Best Local Similarity: 84.1%    Mismatches: 5
  Query Match:      75.0%        Indels:      49
  DB:               4           Gaps:         1

US-10-825-632-4 (1-1197) x US-10-072-012-621 (1-883)

Qy 3 TTTGAAGGACCAAAAGCTCCCTTTAGAGCATCACCTGTACGTAGTCAGTTACGTAAAT 62
Db 525 PheGluGlyThrLysAspSerProLeuGluHisLeuTyrValThrSerTyrAlaAsn 544
Qy 63 CTGGAGAGGTGACAGGCTGACTGACCGTGGCTACTCACATCTTGTGCTGATCAGTCAG 122
Db 545 ProGlyGluValValArgLeuThrAspArgGlyTyrSerHisSerCysLeuSerArg 564
Qy 123 CACTGTGACTCTTTTATAAGTAAGTATAGTAACAGAGAAATCCACACTGTGTGCCCTT 182
Db 565 HisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeu 584
Qy 183 TACAAGCTATCAAGTCCTGAAGATGACCCAACTTGCAGAAACAAAGAAATTTGGGCCACC 242
Db 585 TyrLysLeuSerSerProGluAspAspProValHisLysThrLysGluPheTrpAlaThr 604
Qy 243 ATTTTGGATTTCAGCAGGCTCTCTCTGACTATATCTCTCCAGAAATTTTCTTTTGAA 302
Db 605 IleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePheSerPheGlu 624
Qy 303 AGTACTACTGGATTGATGGGATGCTCTACAGGCTCATGATCTACAGCCTGGA 362
Db 625 SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly 644
Qy 363 AAGAAATATCTTACTGTGTCTTCATATATGTTGGTCTCTCAG----- 404
Db 645 LysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeuValAsnAsn 664
Qy 404 ----- 404
Db 665 ArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrValVal 684
Qy 404 ----- 404
Db 685 ValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLys 704
Qy 405 -----GGTCAATAGAAATTCAGCATCAGGTGGAAGGACTCCCAATATCTAGCTTCT 455
Db 705 TyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyrLeuAlaSer 724
Qy 456 CGATATGATTTCATTGACTTATAGATCGTGTGGGCATCCACGGCTGGTCTTATGAGGATAC 515
Db 725 GlnTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyrGlyGlyTyr 744
Qy 516 CTCTCCCTGATGGCATTAATGCAGAGGTTCAGATATCTTCAGGGTGTCTATTGTGGGGCC 575
Db 745 LeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAlaGlyAla 764
Qy 576 CCAGTCACTCTGTGGATCTTCATGATACAGGATACAGGACGCTTATATGGGTCCACCT 635
Db 765 ProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMetGlyHisPro 784
Qy 636 GACCAGAATGAACAGGGCTATTACTTAGGATCTGTGGCCATCTGAGCAGAGAAATGTCGCC 695
Db 785 AspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGluLysPhePro 804
```

```
Qy 696 TGTGAACCAAAATCGTTTACTCTCTTACATGGTTTCTCGATGAGAAATGTCCATTTTGCA 755
Db 805 SerGluProAsnArgLeuLeuHisGlyPheLeuAspGluAsnValHisPheAla 824
Qy 756 CATACAGTATATTACTGAGTTTTTTAGTGGGCTGAAAGGCATATGATTTACAGATC 815
Db 825 HisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAspLeuGlnIle 844
Qy 816 TATCTCTCAGGAGACACAGCATAGAGTTCCTGAATCGGAGAACATTTATGAAGTCAT 875
Db 845 TyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyrGluLeuHis 864
Qy 876 CTTTTCACCTACCTTCAAGAAAACCTTGGATCAGCTATTGCTCTTAAAGTGATA 932
Db 865 LeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaLeuLysValIle 883
```

RESULT 12

```
US-10-825-632-5
; Sequence 5: Application US/10825632
; Publication No: US20040191826A1
; GENERAL INFORMATION:
; APPLICANT: ABBOTT, Catherine Anne
; APPLICANT: CORRELL, Mark Douglas
; TITLE OF INVENTION: DYPEPTIDYL PEPTIDASES
; FILE REFERENCE: FCSB-100-Div. 1
; CURRENT APPLICATION NUMBER: US/10/825,632
; CURRENT FILING DATE: 2004-04-15
; PRIOR APPLICATION NUMBER: US 10/070,464
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: PCT/AU00/01085
; PRIOR FILING DATE: 2000-09-11
; PRIOR APPLICATION NUMBER: AU PQ5709
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: AU PQ2762
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 465
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-825-632-5
```

```
Alignment Scores:
Pred. No.:      7,76e-137      Length:      465
Score:          1377.50        Matches:      259
Percent Similarity: 83.5%      Conservative: 0
Best Local Similarity: 83.5%    Mismatches: 0
Query Match:      64.4%        Indels:      51
DB:               4           Gaps:         1
```

US-10-825-632-4 (1-1197) x US-10-825-632-5 (1-465)

```
Qy 3 TTTGAAGGACCAAAAGCTCCCTTTAGAGCATCACCTGTACGTAGTCAGTTACGTAAAT 62
Db 207 PheGluGlyThrLysAspSerProLeuGluHisLeuTyrValSerTyrValAsn 226
Qy 63 CTGGAGAGGTGACAGGCTGACTGACCGTGGCTACTCACATCTTCTGCTCAGTCAGTCAG 122
Db 227 ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysValSerGln 246
Qy 123 CACTGTGACTCTTTTATAAGTAAGTATAGTAACAGAGAAATCCACACTGTGTGCCCTT 182
Db 247 HisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeu 266
Qy 183 TACAAGCTATCAAGTCTCTGAAATGACCCAACTTGCAGAAACAAAGAAATTTGGGCCACC 242
Db 267 TyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPheTrpAlaThr 286
Qy 243 ATTTTGGATTTCAGCAGGCTCTCTCTGACTATATCTCTCAGAAATTTTCTTTTGAA 302
Db 287 IleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePheSerPheGlu 306
```



```
QY 303 AGTACTACTGGAATTACATTTGATGGGATGCTCTACAGCCTCATGATCTACAGCCTGGA 362
Db 307 SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly 326
QY 363 AAGAAATACCTACTGCTGCTCATATATGCTGCTCTACAGGCTCAATAGAAATTGAC 422
Db 327 LysLysTyrProThrValLeuPheIleTyrGlyGlyProGln----- 340
QY 423 GATCAGGTGGAAGGACCTCAATATCTAGCTTCTCGATGATGATTTTCATTGCACTTAGATCGT 482
Db 340 ----- 340
QY 483 GTGGGCATCCACGGCTGCTCTATGGAGGATACCTCTCCTGATGGCATTTAATGCAGAGG 542
Db 340 ----- 340
QY 543 TCAGATATCTCAGGGTGTCTATTGCTGGGGCCCGACCTCCTCTGTCGATCTTCTATGAT 602
Db 341 -----ValAlaIleAlaGlyAlaProValThrLeuTrpIlePheTyrAsp 355
QY 603 ACAGGATACACGGAACGTTATATGGGTCACTGACCAAGATGAACAGGCTATTACTTGA 662
Db 356 ThrGlyTyrThrGluArgTyrMetGlyHisProAspGlnAsnGluGlnGlyTyrTyrLeu 375
QY 663 GdATCTGTGGCCATCGAAGCAGAAAGTTCCCTCTGAAACCAAAATCGTTTACTGCTCTTA 722
Db 376 GlySerValAlaMetGlnAlaGluLysPheProSerGluProAsnArgLeuLeuLeu 395
QY 723 CATGGTTTCTGGATGAGATGTCATTTTGGCATACACAGTATATATTACTGAGTTTTTTA 782
Db 396 HisGlyPheLeuAspGluAsnValHisPheAlaHisThrSerIleLeuLeuSerPheLeu 415
QY 783 GTGAGGGCTGGAAGCCATATGATTTTACAGATCTATCTCTCAGGAGACACAGCATAGA 842
Db 416 ValArgAlaGlyLysProTyrAspLeuGlnIleTyrProGlnGluArgHisSerIleArg 435
QY 843 GTTCTGAATCGGAGAACATTATGAATCGCATCTTTTGCACCTTCTTCAAGAAAAACCTT 902
Db 436 ValProGluSerGlyGluHisTyrGluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeu 455
QY 903 GGATCAGCTATTGCTGCTCTAAAAGTGATA 932
Db 456 GlySerArgIleAlaAlaLeuLysValIle 465
```

RESULT 13

```
US-10-275-505-2
; Sequence 2, Application US/10275505
; Publication No. US20040081961A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: DELEGANE, Angelo M.; LAL, Preeti G.
; APPLICANT: HAFALIA, April J.A.; PATTERSON, Chandra
; APPLICANT: WALIA, Narinder K.; KEARNEY, Liam
; APPLICANT: TRIBOULEY, Catherine M.; KHAN, Farrah A.
; APPLICANT: YAO, Monique G.; BAUGHN, Mariah R.
; APPLICANT: AZIMZAI, Valda; ELLIOTT, Vicki S.
; APPLICANT: NGUYEN, Dannel B.; GANDHI, Ameena R.
; APPLICANT: YANG, Juming; HERNANDEZ, Roberto
; APPLICANT: POLICKY, Jennifer L.; LU, Dyung Aina M.
; APPLICANT: REDDY, Roopa M.; YUE, Henry
; APPLICANT: TANG, Y. Tom
; TITLE OF INVENTION: PROTEASES
; FILE REFERENCE: PI-0085 USN
; CURRENT APPLICATION NUMBER: US/10/275,505
; CURRENT FILING DATE: 2002-11-04
; PRIOR APPLICATION NUMBER: PCT/US01/14651
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: 60/209,402
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: 60/207,477
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/205,803
; PRIOR FILING DATE: 2000-05-17
```

```
; PRIOR APPLICATION NUMBER: 60/203,566
; PRIOR FILING DATE: 2000-05-11
; PRIOR APPLICATION NUMBER: 60/202,082
; PRIOR FILING DATE: 2000-05-04
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PERL Program
; SEQ ID NO 2
; LENGTH: 580
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 376067CDI
US-10-275-505-2
```

Alignment Scores:

```
Pred. No.: 3,22e-136 Length: 580
Score: 1372.00 Matches: 266
Percent Similarity: 87.1% Conservative: 11
Best Local Similarity: 83.6% Mismatches: 19
Query Match: 64.1% Indels: 22
DB: 4 Gaps: 6
```

US-10-825-632-4 (1-1197) x US-10-275-505-2 (1-580)

```
QY 3 TTTGAAGCCACCAAGACTCCCTTTAGAGCATCACCTGTACGTAGTCAGTTACGTAAAT 62
Db 277 PheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSerTyrValAsn 296
QY 63 CCTGGAGAGTGACAGGCTGACTGACCGGTGGCTACTACATCTTCTGTCGATCAGTCAG 122
Db 297 ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCysIleSerGln 316
QY 123 CACTGTCACTCTTTATAGTAAGTATAGTAACCAAGAAATCCACACTGTGTGTCCCTT 182
Db 317 HisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeu 336
QY 183 TACAAGCTATCAAGTCTCTGAAGATGACCAACTTGCACAAACAAAGAAATTTTGGGCCACC 242
Db 337 TyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPheTrpAlaThr 356
QY 243 ATTTTGAATTCAGAGGCTCTCTCTGACTATATCTCTCCAGAAATTTTCTTTTGA 302
Db 357 IleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePheSerPheGlu 376
QY 303 AGTACTACTCGATTACATTTGATGGATGCTCTACAGCCTCATGATCTACAGCCTGGA 362
Db 377 SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly 396
QY 363 AAGAAATATCTACTGTGCTGTTTCATATATGTTGTTGCTCTCAGGGTCAATAAGAAATTGAC 422
Db 397 LysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu---ValAsn 415
QY 423 GATCAGGTGGAAGGACTCCCAATAT-----CTAGCTTCT---CGATATGAT 464
Db 416 AsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrVal 435
QY 465 TTCATTGACTTAGATCGTGTGGGCATC-----CAGGCTGCTCTATGGAGGATACCTC 518
Db 436 ValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPhe 455
QY 519 TCCCTGATGCATTAAATGCAGAGGTCAATATCTTCCAGGGTGTCTTATTTGCTGGGGCCCCA 578
Db 456 LysTyrLys-----MetValAlaIleAlaGlyAlaPro 466
QY 579 GTCACTCTGTGGATCTTCTATGATACAGGATACACGGAACGTTATATGGGTCACTCCCTGAC 638
Db 467 ValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMetGlyHisProAsp 486
QY 639 CAGAAATGAACAGGGCTATTACTTAGGATCTGTGGCATGCAAGCAGAAAGTTCCCTCT 698
Db 487 GlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGluLysPheProSer 506
```

699	QY	GAACCAAAATCGTTTACTGCTCTTACATGCTTTCTCGATGAGAAATGCTCCATTTTGCACAT	758
507	Db	GLIPRoAsnArgLeuLeuLeuLeuHiGlyPheLeuAspGluAsnValHisPheAlaHis	526
759	QY	ACCAGTATATTACTGAGTGTATTTTATGAGGGCTGGAAAGCCATATGATTTACAGATCTAT	818
527	Db	ThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAspLeu-----	543
819	QY	CCTCAGGAGAGACACAGCATAAGAGTTCCTGAATCGGGAGAACATATTATGAATCGCATCTT	878
544	Db	---GlnGluArgHisSerIleArgValProGluSerGlyGluHisTyrGluLeuHisLeu	562
879	QY	TTGCACCTACCTTCAAGAAAACCTTCGGATCAGCTATTGCTGCTCTAAAAAGTCATA	932
563	Db	LeuHisTyrIleuGlnGluAsnLeuGlySerArgIleAlaLeuLeuValIle	580

Db	277	PhcGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSerTyrValAsn	296
Qy	63	CCTGGAGAGGTGACAAAGCTGACGCGGTGCTACTCACATCTCTGCTGCATCAGTCAG	122
Db	297	ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysValleSerGln	316
Qy	123	CACTGTGACTCTTTTATAGTATAGTATAGTAAACAGAGAAGATCCACACTGTGTGTCCTT	182
Db	317	HisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeu	336
Qy	183	TACAAGCTATCAAGTCTCTGAAGATGACCCAACTTGGCAAAAACAAGGAATTTTGGGCCACC	242
Db	337	TyLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPheTrpAlaThr	356
Qy	243	ATTTTGGATTTCAGCAGGTCCTCTTCTGACTATATCTCCTCCAGAAATTTTCTCTTTTGA	302
Db	357	IleLeuAspSerAlaGlyProLeuProAspTyrThrProGluLlePheSerPheGlu	376
Qy	303	AGTACTACTGGATTTTACATCTGTATGGGATGCTCTACAAAGCCTCATGATCTACAGCCTGGA	362
Db	377	SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly	396
Qy	363	AGAAATATCTTACAGTGTGCTGTTCATATATCGTGGTCTCTCAGGGTCAAATAGAAAATGAC	422
Db	397	LysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeuValAsn	415
Qy	423	GATCAGGTGGAAGGACTCCAATAT-----CTAGCTCTCT--CGATATGAT	464
Db	416	AsnArgPheGlyGlyValIysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrVal	435
Qy	465	TTCAITGACTTAGATCGTGTGGGCATC-----CACGGCTGGCTTATGGAGGATACCTC	518
Db	436	valValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPhe	455
Qy	519	TCCTCGATGGCATTAAATCAGAGGTTCAGATATCTTCAGGGTTGCTATTGCTGGGGCCCCA	578
Db	456	LysTyrLys-----MetValAlaIleAlaGlyAlaPro	466
Qy	579	GTCACTCTGTGATCTTCTATGATACAGGATACAGGAAGCTTATATGGGTACCCCTGAC	638
Db	467	ValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMetGlyHisProAsp	486
Qy	639	CAGAATGAACAGGGCTATTACTTTAGGATCTCTGGCCATGCAAGCAGAAAAGTTCCCTCT	698
Db	487	GlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGluLysPheProSer	506
Qy	699	GAACAAATCTGTTACTGCTCTTACATCGTTTCTTGATGAGATGTCATTTTGCACAT	758
Db	507	GluProAsnArgLeuLeuLeuLeuHisGlyPheLeuLeuAspGluAsnValHisPheAlaHis	526
Qy	759	ACCAATATATTACTGAGTTTTTACTGAGGGCTGAAAGCCATATGATTTTACAGATCTAT	818
Db	527	ThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAspLeu-----	543
Qy	819	CCTCAGGAGACACACAGATAGAGTCTCTGATCGGAGACACATATGACTGCATCTT	878
Db	544	--GlnGluArgHisSerIleArgValProGluSerGlyGluHisTyrGluLeuHisLeu	566
Qy	879	TTCGCTACTCTTCAAGAAAACCTTGATCACGATTTGCTGCTCTTAAAAGTGATA	932
Db	563	LeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeuLysValIle	580
RESULT 15			
US-10-415-122-4			
; Sequence 4, Application US/10415122			
; Publication No. US20040053369A1			
; GENERAL INFORMATION:			
; APPLICANT: THE UNIVERSITY OF SYDNEY			
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES			
; FILE REFERENCE: PP15217			
; CURRENT APPLICATION NUMBER: US/10/415,122			
; CURRENT FILING DATE: 2003-08-07			

```
; NUMBER OF SEQ-IDS NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 869
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-415-122-4

Alignment Scores:
Pred. No.: 1,13e-106      Length: 869
Score: 1094.50           Matches: 200
Percent Similarity: 69.3% Conservative: 42
Best Local Similarity: 57.3% Mismatches: 58
Query Match: 51.1%      Indels: 49
DB: 4                    Gaps: 1

US-10-825-632-4 (1-1197) x US-10-415-122-4 (1-869)

QY 3 TTTGAAGGACCAAGACTCCCTTTAGAGCATCAGCTGTAGTAGTCAGTTACGTAAAT 62
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 62
QY 521 PheGlnGlyThrLysAspThrProLeuGluHisHisLeuTyrValValSerTyrGluSer 540
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||||| 540
QY 63 CCTGGAGAGGTGACAGGCTGACTGACCGTGCTACTCACATTTCTTGTGTCAGTCAGTCAG 122
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||||| 122
QY 541 AlaGlyGluLeuValArgLeuThrThrLeuGlyPheSerHisSerCysSerMetSerGln 560
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||||| 560
QY 123 CACTGTGACTCTTTTAAAGTAAAGTATAGTAAACAGAGAATCCACACTGTGTGTCCTT 182
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||||| 182
QY 561 SerPheAspMetPheValSerHisTyrSerSerValSerThrProProCysValHisVal 580
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||||| 580
QY 183 TACAGCTATCAAGTCTCTGAGATGACCCCACTTCGAAACAAAGGAATTTTGGGCCACC 242
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||||| 242
QY 581 TyrLysLeuSerGlyProAspAspProLeuHisLysGlnProArgPheTrpAlaSer 600
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||||| 600
QY 243 ATTTTGGATTACAGCAGGTCTCTTCTGACTACTACTCTCCAGAAATTTCTCTTTTGA 302
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||||| 302
QY 601 MetMetGluAlaAlaAsnCysProProAspTyrValProProGluLeuPheHisPheHis 620
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||||| 620
QY 303 AGTACTACTGGATTTACATTTGATGGGATGCTCTCAAGCCTCATGATCTACAGCCTGGA 362
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||||| 362
QY 621 ThrArgAlaAspValGlnLeuTyrGlyMetIleTyrLysProHisThrLeuGlnProGly 640
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||||| 640
QY 363 AAGAAATATCTACTGTGCTGTTCATATATGTGTGCTCTCAG----- 404
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||||| 404
QY 641 ArgLysHisProThrValLeuPheValTyrGlyGlyProGlnValGlnLeuValAsnAsn 660
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||||| 660
QY 404 ----- 404
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||||| 404
QY 661 SerPheLysGlyIleLysTyrLeuArgLeuAsnThrLeuAlaSerLeuGlyTyrAlaVal 680
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||||| 680
QY 404 ----- 404
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||||| 404
QY 681 ValValIleAspGlyArgGlySerCysGlnArgGlyLeuHisPheGluGlyAlaLeuLys 700
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||||| 700
QY 405 -----GGTCAATAGAAATTGACGATCAGGTGGAGGACTCCCAATATCTAGCTTCT 455
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||||| 455
QY 701 AsnGlnMetGlyGlnValGluLeuAspGlnValGluGlyLeuGlnTyrValAlaGlu 720
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||||| 720
QY 456 CGATATGATTTCTATTCATCTAGATCTGTGGGCATCCACGGCTGCTCTATGGAGGATAC 515
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||||| 515
QY 721 LysTyrGlyPheIleAspLeuSerArgValAlaIleHisGlyIleTrpSerTyrGlyPhe 740
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||||| 740
QY 516 CTCTCCCTGATGCGATTAAATGACAGGTTCAGATATCTTCAGGGTTGCTATTGCTGGGGCC 575
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||||| 575
QY 741 LeuSerLeuMetGlyLeuIleHisLysProGlnValPheLysValAlaIleAlaGlyAla 760
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||||| 760
QY 576 CCACTCATCTGTGGATCTTCTATGATACAGGATACACGGAACGTATATGGGTCACCT 635
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||||| 635
QY 761 ProValThrValTrpMetAlaTyrAspThrGlyTyrThrGluArgTyrMetAspValPro 780
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||||| 780
QY 636 GACCAGATGACAGGGCTATTACTAGTCTGTGGCCATGCAAGCAGAAAGTCCCC 695
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||||| 695
QY 781 GluAsnAsnGlnGlyTyrGluAlaGlySerValAlaLeuHisValGluLysLeuPro 800
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||||| 800
```

Search completed: May 2, 2006, 03:08:34
Job time : 190.411 secs

```
QY 696 TCTGAACCAAAATCGTTTACTGCTCTTACATGGTTTCTCGATGAGAAATGTCATTTCGA 755
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||||| 755
QY 801 AsnGluProAsnArgLeuLeuLeuLeuHisGlyPheLeuAspGluAsnValHisPhePhe 820
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||||| 820
QY 756 CATACCACTATATTAAGTGTGTTTGTAGTGGGCTGGAAGCCCATATGATTTACAGATC 815
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||||| 815
QY 821 HisThrAsnPheLeuValSerGlnLeuIleArgAlaGlyLysProTyrGlnLeuGlnIle 840
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||||| 840
QY 816 TATCTCTCAGGAGAGACACAGACATAGAAGATTCTCTGAATCGGAGAACATTATGAATGCGAT 875
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||||| 875
QY 841 TyrProAsnGluArgHisSerIleArgCysArgGluSerGlyGluHisTyrGluValThr 860
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||||| 860
QY 876 CTTTGTGACACTACCTTCAAGAAAACCTT 902
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||||| 902
QY 861 LeuLeuHisPheLeuGlnGluHisLeu 869
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.1.7
Copyright (c) 1993 - 2006 Biocolloration Ltd.

OM nucleic - protein search using (same plus n2p model)

Run on: May 2, 2006, 00:41:48 ; Search time 36.3435 Seconds
(without alignments)
2894.257 Million cell updates/sec

Title: US-10-825-632-4
Perfect score: 2140
Sequence: 1 attttgaaggcaccacaaagac.....ttactaataaaaaaaaaa 1197

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2443163 seqs, 439378781 residues
Total number of hits satisfying chosen parameters: 4886326

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ n2p.model -DEV=xlp
-O=/abs/ABSSWEB.spool/US10825632/runat_01052006.105944.3214/app query.fasta.1
-DB=A Geneseq -QFWT=fastcan -SUFFIX=n2p.rag -MINMATCH=0.1 -LOOFCI=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFWT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abs02p
-USER=US10825632 @CGN 1.1.605 @runat_01052006.105944.3214 -NCPU=6 -ICPU=3
-NO MAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A Geneseq 21.*

- 1: geneseqp1980s.*
- 2: geneseqp1990s.*
- 3: geneseqp2000s.*
- 4: geneseqp2001s.*
- 5: geneseqp2002s.*
- 6: geneseqp2003as.*
- 7: geneseqp2003bs.*
- 8: geneseqp2004s.*
- 9: geneseqp2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1680	78.5	310	4	AAB47188 Human DPP
2	1680	78.5	310	5	AB080994 Human dip
3	1680	78.5	310	7	AD27858 Human dip
4	1645.5	76.9	587	5	ADR41398 Human CD-
5	1645.5	76.9	824	6	ABU92030 Human pro
6	1645.5	76.9	882	4	AAB47187 Human DPP
7	1645.5	76.9	882	5	ABG61591 Human DPP
8	1645.5	76.9	882	5	AAG78415 Amino aci
9	1645.5	76.9	882	5	AAG24170 Human dip

10	1645.5	76.9	882	5	AAU74749	Aau74749 Human pro
11	1645.5	76.9	882	5	ADI17086	Adi17086 Human NOV
12	1645.5	76.9	882	6	ABU07720	Abu07720 Human ser
13	1605.5	75.0	883	5	ADI17085	Adi17085 Murine NO
14	1407	65.7	493	7	ADE78977	Ade78977 Human pro
15	1407	65.7	831	6	ABU92026	Abu92026 Human pro
16	1401	65.5	632	4	ABU93565	Abu93565 Human pro
17	1377.5	64.4	465	4	AAB47189	Aab47189 Human DPP
18	1377.5	64.4	724	5	ABU97362	Abu97362 Novel hum
19	1377.5	64.4	782	5	ABB97361	Abb97361 Novel hum
20	1372	64.1	516	6	ABU92029	Abu92029 Human pro
21	1372	64.1	580	5	AAE14337	Aae14337 Human pro
22	1337	62.5	738	8	ADT04044	Adt04044 Human pro
23	1094.5	51.1	869	5	AAE24169	Aae24169 Alternati
24	1093.5	51.1	497	5	AAU96169	Aau96169 Human sec
25	1093.5	51.1	497	5	ABG64845	Abg64845 Human alb
26	1093.5	51.1	497	8	ADL78112	Adl78112 Albumin f
27	1093.5	51.1	597	7	ADE08485	Ade08485 Novel pro
28	1093.5	51.1	689	5	ADI17084	Adi17084 Human NOV
29	1093.5	51.1	830	5	AAE24171	Aae24171 Human dip
30	1093.5	51.1	863	5	ABG61592	Abg61592 Human DPP
31	1093.5	51.1	863	5	ADI17083	Adi17083 Human NOV
32	1093.5	51.1	892	5	ABG61602	Abg61602 Human DPR
33	1093.5	51.1	892	5	ABG61604	Abg61604 Human DPR
34	1093.5	51.1	892	5	ABB98134	Abb98134 Human PMW
35	1093.5	51.1	892	8	ADS10951	Ads10951 Human the
36	1093.5	51.1	969	5	AAE24168	Aae24168 Human dip
37	1089.5	50.9	497	5	AAU96192	Aau96192 Human sec
38	1089.5	50.9	497	5	ABG64844	Abg64844 Human alb
39	1089.5	50.9	497	8	ADL78111	Adl78111 Albumin f
40	1089.5	50.9	581	5	ADR41521	Adr41521 Human CD-
41	1086.5	50.8	360	4	AAB47190	Aab47190 Human DPP
42	1086.5	50.8	755	8	ADQ67811	Adq67811 Novel hum
43	1086.5	50.8	879	5	ABG61607	Abg61607 Human DPR
44	1086.5	50.8	879	5	ABG61608	Abg61608 Human DPR
45	1086.5	50.8	879	8	ADS10953	Ads10953 Human the

ALIGNMENTS

RESULT 1
AAB47188
ID AAB47188 standard; protein; 310 AA.
XX
AC AAB47188;
AC AAB47188;
DT 29-JUN-2001 (first entry)
XX
DE Human DPP8 524Phe-657Gln+707Gly-882Ile.
XX

Human; dipeptidyl aminopeptidase; DPP8; prolly oligopeptidase;
KW dipeptidyl peptidase; DPPIV; T cell; cleavage; diarrhoea;
KW growth hormone deficiency; glucose level; mucosal regeneration;
KW non-insulin dependent diabetes mellitus; glucose intolerance;
KW immunosuppression.

XX Homo sapiens.
XX
XX WO200119866-A1.
XX
XX 22-MAR-2001.
XX

XX 11-SEP-2000; 2000WO-AU001085.
XX
XX 10-SEP-1999; 99AU-00002762.
XX 18-FEB-2000; 2000AU-00005709.
XX (UNSY) UNIV SYDNEY.
XX
XX Abbott CA, Gorell MD;
XX
XX WPI; 2001-281520/29.
XX
XX N-PSDB; AAC85695.

XX New human dipeptidyl aminopeptidase (DPP8) useful for cleaving
PT substrates, identifying inhibitors of DPP8 catalytic activity which have
PT therapeutic uses, and for detecting activated T cells.

XX Claim 5; Page 68-69; 78pp; English.

XX The sequences given in ABA47188-90 represent fragments of human
CC dipeptidyl aminopeptidase (DPP8). DPP8 has substrate specificity for H-
CC Ala-Pro-pNA, H-Gly-Pro-pNA and H-Arg-Pro-pNA. Therefore, it is a prolyl
CC oligopeptidase and a dipeptidyl peptidase, because it is capable of
CC hydrolysing the peptide bond C-terminal to pro in each of these
CC compounds. DPP8 is homologous with human DPPIV. DPP8 is useful for
CC cleaving a substrate, and for detecting an activated T cell which
CC involves measuring the level of DPP8 gene expression in a T cell. The
CC level of DPP8 expression is detected by identifying a molecule capable of
CC inhibiting the cleavage of the substrate by DPP8. Molecules identified as
CC inhibiting DPP8 catalytic activity may be useful for treating diarrhoea,
CC growth hormone deficiency, lowering glucose levels in non-insulin
CC dependent diabetes mellitus and other disorders involving glucose
CC intolerance, enhancing mucosal regeneration and as immunosuppressants
XX Sequence 310 AA;

Alignment Scores:

Pred. No.: 1 886-184 Length: 310
Score: 1680.00 Matches: 310
Percent Similarity: 100.0% Conservatives: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 78.5% Indels: 0
DB: 4 Gaps: 0

US-10-825-632-4 (1-1197) x ABA47188 (1-310)

Qy 3 TTGAAGGACCAAGAGCTCCCTTTAGAGATCACCTGTACGTCAGTCAGTTACGTAAT 62
Db 1 PheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSerTyrValAsn 20
Qy 63 CCTGGAGAGTGACAGGCTGACCTGACCTGCTACTACATCTTCTGTCATCAGTCAG 122
Db 21 ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysIleSerGln 40
Qy 123 CACTGTGACTCTTTTATAAGTATAGTAAACAGAGAGATCCACACTGTGTGCCCTT 182
Db 41 HisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeu 60
Qy 183 TACAAGCTATCAAGTCTCAAGATGACCACTTGCAAAACAAAGGAATTTTGGGCCACC 242
Db 61 TyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPheTyrAlaThr 80
Qy 243 ATTTTGATTTCAGCAGGTCCTCTCTGACTATATCTCTCCAGAAATTTTCTCTTTGAA 302
Db 81 IleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePheSerPheGlu 100
Qy 303 AGTACTACTGGATTTACATGTTATGGATGCTCTACAGCCTCATGATCTACAGCCTGGA 362
Db 101 SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly 120
Qy 363 AAGAAATACCTACTGTCTGTCTATATATGCTGCTCAGGTCGAATAGAAATGAC 422
Db 121 LysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnGlyGlnIleGluIleAsp 140
Qy 423 GATCAGGTGAAGGACTCCAATATCTAGCTTCGATATGATTTTCATTGATTCATCGT 482
Db 141 AspGlnValGluGlyLeuGlnTyrLeuAlaSerArgTyrAspPheIleAspLeuAspArg 160
Qy 483 GTGGGCATCACCGCTGGTCTTATGGAGGATACCTCTCCCTGATGGCATTAATGCGAGG 542
Db 161 ValGlyIleHisGlyTrpSerTyrGlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArg 180
Qy 543 TCAGATATCTTCAGGTTCTCTATGCTGGGCCCCAGTCACCTCTGTGATCTTCTATGAT 602

Db 181 SerAspIlePheArgValAlaIleAlaGlyAlaProValThrLeuTyrIlePheTyrAsp 200
Qy 603 ACAGGATACACGAGACCTTATATGGTCCACCTGACACAGATGACAGGCTATTACTTA 662
Db 201 ThrGlyTyrThrGluArgTyrMetGlyHisProAspGlnAsnGlnGlyTyrTyrLeu 220
Qy 663 GGATCTGTGGCCATGACAGCAAGAAAGTCCCTCTGAAACCAATCGTTTACTGCTTTA 722
Db 221 GlySerValAlaMetGlnAlaGluLysPheProSerGluProAsnArgLeuLeuLeu 240
Qy 723 CATGGTTTCTGGATGAGAAATCTCCATTTTCCACATACCATATATTACTGAGTTTATA 782
Db 241 HisGlyPheLeuAspGluAsnValHisPheAlaHisThrSerIleLeuLeuSerPheLeu 260
Qy 783 GTGAGGCTGGAAAGCCATATGATTTACAGATCTATCTCAGGAGACACAGCATAGA 842
Db 261 ValArgAlaGlyLysProTyrAspLeuGlnIleTyrProGlnGluArgHisSerIleArg 280
Qy 843 GTTCTCTGAATCGGGAGAACATTTATGAATGTCATCTTTTGCACACTACCTTCAAGAAAACCTT 902
Db 281 ValProGluSerGlyGluHisTyrGluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeu 300
Qy 903 GGATCAGTATTGCTGCTCTTAAAGTGATA 932
Db 301 GlySerArgIleAlaLeuLysValIle 310

RESULT 2

ABB08994
ID ABB08994 standard; protein; 310 AA.
XX
AC ABB08994;
XX
DT 19-JUN-2002 (first entry)
XX
DE Human dipeptidyl peptidase 8.
KW Human; dipeptidyl peptidase 8; antiasthmatic; antiallergic;
KW antiinflammatory.
XX
OS Homo sapiens.
XX
FN US6337069-B1.
XX
PD 08-JAN-2002.
XX
PF 28-FEB-2001; 2001US-00794236.
XX
PR 28-FEB-2001; 2001US-00794236.
XX
PA (BMRA-) BMRA CORP BV.
XX
PI Grouzmann E, Lacroix J, Monod M;
XX
DR WPI; 2002-163235/21.
XX
PT Treating a patient for mucosal inflammation associated with rhinitis,
PT sinusitis or both, by intranasally administering a peptidase that cleaves
PT at Xaa-Pro sequences, to the patient.
XX
PS Disclosure; Col 23-24; 13pp; English.
XX
CC Thus invention relates to the treating of a patient for mucosal
CC inflammation associated with rhinitis or sinusitis, comprising
CC intranasally administering a peptidase. The peptidase is considered
CC antiasthmatic, antiallergic and antiinflammatory in its action. The
CC peptidase cleaves at Xaa-Pro sequences and is useful for treating a
CC patient for mucosal inflammation associated with rhinitis or sinusitis,
CC which is the result of allergies or asthma. This sequence represents
CC human dipeptidyl peptidase 8
SQ Sequence 310 AA;

Alignment Scores:

Pred. No.:	1.88e-184	Length:	310
Score:	1680.00	Matches:	310
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	78.5%	Indels:	0
DB:	5	Gaps:	0
US-10-825-632-4 (1-1197) x ABB08994 (1-310)			
Qy	3	TTTGAAGGCAACAAAGACTCCCTTTAGAGCATACCTGTAGTACGTAGTCAGTTACGTAAAT	62
Db	1	PheGluGlyThrIysAspSerProLeuGluHisLeuTyrValValSerTyrValAsn	20
Qy	63	CCTGGAGAGTGCACAGGCTGACTGACCTGGCTACTCACATCTTCTGCTGATCAGTCAG	122
Db	21	ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCysIleSerGln	40
Qy	123	CACGTGACATCTTTTATAGTATAGTAACTACACAGAAATCCACACTGTGTCTCCCTT	182
Db	41	HisCysAspPhePheIleSerIysTyrSerAsnGlnLysAsnProHisCysValSerLeu	60
Qy	183	TACAAGCTATCAAGTCTCTGAAGATGACCCCACTTGCACAAACAAAGGAATTTTGGCCACC	242
Db	61	TyrlsLeuSerSerProGluAspAspProThrCysLysThrLysGluPheTrpAlaThr	80
Qy	243	ATTTTGATTCACAGGTCTCTCTGACTACTACTCTCCAGAAATTTCTCTTTTGA	302
Db	81	IleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePheSerPheGlu	100
Qy	303	AGTACTACTGGATTTACATTGATGGGATGCTCTCAAGCCTCATGATCTACAGCTGGA	362
Db	101	SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly	120
Qy	363	AAGAAATATCTACTGTCTTCATATATGTGTCTCTCAGGTCAAAATAGAAATTCAC	422
Db	121	LysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnGlyGlnIleGluIleAsp	140
Qy	423	GATCAGTGGAGGACTCCAAATATCTAGCTCTCCATATGATTTTCATTGACTAGATCGT	482
Db	141	AspGlnValGluGlyLeuGlnTyrLeuAlaSerArgTyrAspPheIleAspLeuAspArg	160
Qy	483	GTGGGATCCACGGCTGGTCTTATGGAGGATACCTCTCCCTGATGGCATTAAATCGAGG	542
Db	161	ValGlyIleHisGlyTrpSerTyrGlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArg	180
Qy	543	TCAGATATCTTCAGGTTGCTATTTGCTGGGCCCCAGTCACCTCTGTGATCTTCTATGAT	602
Db	181	SerAspIlePheArgValAlaIleAlaGlyAlaProValThrLeuTrpIlePheTyrAsp	200
Qy	603	ACGGATACACGGAACGTTATATGGGTACCTCTGACAGAAATGAACAGGGCTATTACTTA	662
Db	201	ThrGlyTyrThrGluArgTyrMetGlyHisProAspGlnAsnGluGlnGlyTyrTyrLeu	220
Qy	663	GGATCTGTGCGCATGCAACGAGAAAGTTCCCTCTGAACCAATCGTTTACTGCTCTTA	722
Db	221	GlySerValAlaMetGlnAlaGluLysPheProSerGluProAsnArgLeuLeuLeu	240
Qy	723	CATGGTTCTCTGATCAGAAATGTCACATTTTGACATACACAGTATATATTACTGAGTTT	782
Db	241	HisGlyPheLeuAspGluAsnValHisPheAlaHisThrSerIleLeuLeuSerPheLeu	260
Qy	783	GTGAGGCTGGAAAGCCATATGATTTTACAGATCTATCTCAGGAGAGACACAGCATAAGA	842
Db	261	ValArgAlaGlyLysProTyrAspLeuGlnIleTyrProGlnGluArgHisSerIleArg	280
Qy	843	GTTCTCTGAATCGGGAGAACATTATGAACATGCTCTTTTGCACACTACCTTCAAGAAACCTT	902
Db	281	ValProGluSerGlyGluHisTyrGluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeu	300
Qy	903	GGATCAGTATTCTCTCTTAAAGTGATA	932
Db	301	GlySerArgIleAlaLeuLysValIle	310

RESULT 3			
ADD27858			
ID	ADD27858 standard; protein; 310 AA.		
XX			
AC	ADD27858;		
XX			
DT	15-JAN-2004 (first entry)		
XX			
DE	Human dipeptidyl peptidase 8.		
XX			
KW	Mucosal inflammation; rhinitis; sinusitis; exopeptidase; substance P; SP; neurokinin 1 receptor; NK1 receptor; allergy; asthma; antiasthmatic; human; dipeptidyl peptidase 8; enzyme.		
KW	antiinflammatory; antiasthmatic; human; dipeptidyl peptidase 8; enzyme.		
XX			
OS	Homo sapiens.		
XX			
PN	US2003165489-A1.		
XX			
PD	04-SEP-2003.		
XX			
PF	27-NOV-2001; 2001US-00993959.		
XX			
PR	28-FEB-2001; 2001US-00794236.		
XX			
PA	(BMEA-) BMEA CORP BV.		
XX			
PI	Grouzmann E, Lacroix J, Monod M;		
XX			
DR	WPI; 2003-811386/76.		
XX			
PT	Treatment of patient for mucosal inflammation associated with rhinitis and/or sinusitis involves intranasally administering peptidase that cleaves at Xaa-Pro sequences or agent inhibiting binding of Sp to neurokinin 1 receptor.		
XX			
PS	Disclosure; SEQ ID NO 4; 14pp; English.		
XX			
CC	The present invention relates to a method of treating a patient for mucosal inflammation associated with rhinitis and/or sinusitis. The method comprises intranasally administering to the patient a peptidase that cleaves at Xaa-Pro sequences or an agent that inhibits the binding of substance P (SP) to the neurokinin 1 (NK1) receptor. The peptidase is an exopeptidase, preferably selected from human dipeptidyl peptidase IV (DPPIV), human quiescent cell proline dipeptidase, human dipeptidyl peptidase 8, or human attractin. The method is useful for treating a patient for mucosal inflammation associated with rhinitis and/or sinusitis which are the result of allergies or asthma. The invention provides an effective treatment of the inflammation associated with both rhinitis and sinusitis. The present sequence represents human dipeptidyl peptidase 8.		
XX			
SQ	Sequence 310 AA;		
Alignment Scores:			
Pred. No.:	1.88e-184		
Score:	1680.00		
Percent Similarity:	100.0%		
Best Local Similarity:	100.0%		
Query Match:	78.5%		
DB:	7		
Length:			
Matches:	310		
Conservative:	0		
Mismatches:	0		
Indels:	0		
Gaps:	0		
US-10-825-632-4 (1-1197) x ADD27858 (1-310)			
QY	3	TTTGAAGGCAACAAAGACTCCCTTTAGAGCATCACCTGTAGTACGTAGTCAGTTACGTAAT	62
Db	1	PheGluGlyThrIysAspSerProLeuGluHisLeuTyrValValSerTyrValAsn	20
QY	63	CCTGGAGAGTGCACAGGCTGACTGACCGCTGGCTACTCACATCTTGTGTCATCAGTCAG	122
Db	21	ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCysIleSerGln	40
QY	123	CACGTGACATCTTTTATAGTAAGTATAGTAACTACACAGAAATCCACACTGTGTGTCCCTT	182

Db 41 HisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeu 60
QY 183 TACAAGCTATCAAGTCTGGAAGATGACCCCAACTTGGCAAAACAAGGAATTTTGGGCACC 242
Db 61 TyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPheTrpAlaThr 80
QY 243 ATTTTGGATTACAGCAGGTCCTCTTCTGACTATATCTCTCCAGAAATTTTCTTTTGA 302
Db 81 IleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePheSerPheGlu 100
QY 303 AGTACTACTGGATTTTACATTTGATGGGATGCTCTACAAAGCTCATGATCTACAGCTCGA 362
Db 101 SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly 120
QY 363 AGCAATATCTACTGCTGCTTCATATATGCTGCTCAGGTCAGGTCAGAAATAGAC 422
Db 121 LysLysTyrProThrValLeuPheIleTyrGlyProGlnGlyGlnIleGluIleAsp 140
QY 423 GATCAGGTGGAAGGACTCAATATCTAGCTTCTCGATATGATTTTCATTTAGATCT 482
Db 141 AspGlnValGluGlyLeuGlnTyrLeuAlaSerArgTyrAspPheIleAspLeuAspArg 160
QY 483 GTGGGATCCACGGCTGCTTCATGAGGATACCTCTCCCTGATGGCATTAATGACAGG 542
Db 161 ValGlyIleHisGlyTrpSerTyrGlyTyrLeuSerLeuMetAlaLeuMetGlnArg 180
QY 543 TCAGATATCTCAGGTTGCTATTGCTGGGCCCCAGTCACTCTGTGGATCTTCTATGAT 602
Db 181 SerAspIlePheArgValAlaIleAlaGlyAlaProValThrLeuTrpIlePheTyrAsp 200
QY 603 ACAGGATACACGGAAGCTTATATGGTCCACCTGACCAAGATGAACAGGCTATTACTTA 662
Db 201 ThrGlyTyrThrGluArgTyrMetGlyHisProAspGlnAsnGluGlnGlyTyrTyrLeu 220
QY 663 GGATCTGTGGCCATGCAAGCAGAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCTTA 722
Db 221 GlySerValAlaMetGlnAlaGluLysPheProSerGluProAsnArgLeuLeuLeu 240
QY 723 CATGTTTCTCGATGAGAATGTCATTTTGCATACCACTGATATATTACTGAGTTTCTTA 782
Db 241 HisGlyPheLeuAspGluAsnValHisPheAlaHisThrSerIleLeuLeuSerPheLeu 260
QY 783 GTGAGGCTGGAAAGCCATGATTTACAGATCTATCTCAGGAGACACAGCATAAGA 842
Db 261 ValArgAlaGlyLysProTyrAspLeuGlnIleTyrProGlnArgHisSerIleArg 280
QY 843 GTTCTGATCGGAGACATTTATCACTGATCTTTTGCACCTTCTTCAAGAAACCTT 902
Db 281 ValProGluSerGlyGluHisTyrGluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeu 300
QY 903 GATCAGCTATTGCTGCTCTAAAGTGATA 932
Db 301 GlySerArgIleAlaAlaLeuLysValIle 310

RESULT 4
ID ADR41398
AC ADR41398 standard; protein; 587 AA.
XX ADR41398;
XX
DT 07-OCT-2004 (first entry)
DE Human CD-like molecule HDQES11, SEQ ID NO:197.
XX
XX Human; CD-like molecule; cluster of differentiation; diagnosis;
KW prevention; immune disorder; immunodeficiency; autoimmune disorder;
KW blood-related disorder; haematological disorder; haemostatic disorder;
KW thrombolytic disorder; hyperproliferative disorder; cancer; tumour;
KW apoptotic disorder; cardiovascular disorder; respiratory disorder;
KW angiogenic disorder; neovascularisation; neurological disorder;
KW endocrine disorder; reproductive system disorder; infectious disease;
KW gastrointestinal disorder; drug screening; tissue regeneration;
KW chemotaxis; gene therapy; antibody therapy; drug targeting;

chromosome mapping; forensic analysis; immunophenotyping; cytostatic;
KW haemostatic; tranquiliser; vulnerary; anti-inflammatory; nephrotropic;
KW cardiant; anti-allergic; anti-HIV; antirheumatic; antiarthritic;
KW antipsoriatic; immunosuppressive; vasotropic; nootropic; neuroprotective;
KW antithyroid; thyrominetic; gynaecological; virucide; hepatotropic;
KW antibacterial; dermatological; chromosome 15q22.
OS Homo sapiens.
XX
XX WO200226930-A2.
XX
PD 04-APR-2002.
XX
PF 25-SEP-2001; 2001WO-US029838.
XX
PR 26-SEP-2000; 2000US-0235484P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Birse CE;
FI WPI; 2002-405050/43.
DR N-PSDB; ADR41222.
XX
PT Novel polynucleotides and polypeptides useful for treating, preventing or
PT ameliorating cardiovascular, renal, neurovascular, and autoimmune
PT disorders.
XX
PS Claim 11; SEQ ID NO 197; 1243pp; English.
XX
XX The invention relates to 167 novel human CD (cluster of differentiation)-
CC like molecules (ADR41388-ADR41563) and to cDNAs encoding them (seqid:1)-
CC
XX Sequence 587 AA;
SQ

Alignment Scores:
Pred. No.: 2,53e-180 Length: 587
Score: 1645.50 Matches: 310
Percent Similarity: 86.4% Conservative: 0
Best Local Similarity: 86.4% Mismatches: 0
Query Match: 76.9% Indels: 49
DB: 5 Gaps: 1

US-10-825-632-4 (1-1197) x ADR41398 (1-587)

QY 3 TTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTAGTACGTAGTACGTAAAT 62
Db 229 PheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSerTyrValAsn 248
QY 63 CCTGGAGAGTGACAGGCTGACTGACCGTGGCTACTCACATTTCTGCTCATCAGTCAG 122
Db 249 ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCysIleSerGln 268
QY 123 CACTGTGACTCTTTTATAGTATAGTATAGTACAGCAGAGAAATCCACACTGTGTCTCCTT 182
Db 269 HisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeu 288
QY 183 TACAGCTATCAAGTCTCTGAAGATGCCCACTTCGCAAAACAAGGAATTTTGGCCACC 242
Db 289 TyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPheTrpAlaThr 308
QY 243 ATTTTGGATTACAGCAGGTCCTCTTCTGACTATATCTCTCCAGAAATTTTCTTTTGA 302
Db 309 IleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePheSerPheGlu 328
QY 303 AGTACTACTGGATTTTACATTTGATGGGATGCTCTCAAGCTCATGATCTACAGCTCGA 362
Db 329 SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly 348
QY 363 AAGAAATATCTACTGCTGCTGTTTCATATATGTTGTTCTCTCTCAG----- 404
Db 349 LysLysTyrProThrValLeuPheIleTyrGlyProGlnValGlnLeuValAsnAsn 368


```
QY 404 ----- 404
Db 369 ArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrValVal 388
QY 404 ----- 404
Db 389 ValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLys 408
QY 405 -----GGTCAAAATAGAAATTGACGTCAGGTGGAAGGACTCCATATCTAGCTTCT 455
Db 409 TyrLysMetGluGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyrLeuAlaSer 428
QY 456 CGATATGATTTTCATTGACTTAGATCGTGTGGGCATCCACGGCTGTCTTATGGAGGATAC 515
Db 429 ArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrPserTyrGlyGlyTyr 448
QY 516 CTCTCCCTGATGGCATTAATGCAGAGTCAGATATCTTCAGGGTTCGTATTGCTGGGGCC 575
Db 449 LeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAlaGlyAla 468
QY 576 CCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAACGTTATATGGTCAACCT 635
Db 469 ProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMetGlyHisPro 488
QY 636 GACCAAGATGAACAGGCTATTACTTAGGATCTGTGGCCATCCAAAGCAGAAAGTTCCCC 695
Db 489 AspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGluLysPhePro 508
QY 696 TCTCAACCAATCGTTTACTGCTTACATGTTCTTGATGTTCTGGATGAGAAATGTCATTGGCA 755
Db 509 SerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnValHisPheAla 528
QY 756 CATACCAAGTATATTACTAGTCTTTTGTAGTGGGCTGGAAGCCATATGATTTACAGATC 815
Db 529 HisthrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAspLeuGlnIle 548
QY 816 TATCTCAGGAGACACACATAAGAGTTCTCGAATCGGAGAGACATTATGAACCTGCAT 875
Db 549 TyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyrGluLeuHis 568
QY 876 CTTTGTGCACTACCTTCAAGAAACCTTGGATCACGTPATTGCTGCTCTAAAGTGATA 932
Db 569 LeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeuLysValIle 587

RESULT 5
ABU92030
ID ABU92030 standard; protein; 824 AA.
AC ABU92030;
XX
XX
XX 15-JUL-2003 (first entry)
XX
XX Human protein modification and maintenance molecule-10 (PMMW-10).
XX
KW Human; protein modification and maintenance molecule; PMMW; cancer;
KW cell proliferation disorder; atherosclerosis; neurological disorder;
KW epilepsy; Huntington's disease; stroke; immune disorder; allergy;
KW inflammatory disorder; AIDS; developmental disorder; hypothyroidism;
KW Cushing's syndrome; gastrointestinal disorder; epithelial disorder;
KW infection; cytostatic; antiarteriosclerotic; anticonvulsant; nootropic;
KW neuroprotective; cerebroprotective; anti-HIV; antiatherogenic;
KW antiinflammatory; thyromimetic.
XX
OS Homo sapiens.
XX
XX WO2003031939-A2.
XX
XX 17-APR-2003.
XX
XX 11-OCT-2002; 2002MO-US032850.
XX
XX 12-OCT-2001; 2001US-0329689P.
XX
XX 25-OCT-2001; 2001US-0335703P.
PR
```

```
PR 09-NOV-2001; 2001US-0348887P.
PR 28-NOV-2001; 2001US-0334145P.
PR 06-DEC-2001; 2001US-0337451P.
PR 14-DEC-2001; 2001US-0340584P.
XX
XX (INCY-) INCYTE GENOMICS INC.
PA
XX
XX RamKumar J, Gorvad AE, Baughn MR, Emerling BM, Yang J, Lee SY,
XX Tran UK, Becha SD, Duggan BM, Lee EA, Griffin JA, Li JX,
XX Sprague WW, Hafalia AUA, Chawla NK, Lehr-Mason PW, Kable AE, Yue H,
XX Marquis JP, Yao MG, Richardson TW, Tang TY, Jin P, Chien D,
XX Bhatia U, Burrill JD, Lee S, Blake JJ, Ho A, Zheng W;
XX WPI: 2003-430274/40.
XX N-PSDB; ACA92425.
XX
XX New human protein modification and maintenance molecules (PMMW), useful
XX for diagnosing, treating and preventing diseases or conditions associated
XX with the aberrant PMMW expression e.g. cancer, atherosclerosis, or
XX infections.
XX
XX Claim 1; Page 249-251; 31pp; English.
XX
XX The present invention relates to the isolation of human protein
XX modification and maintenance molecules (PMMW), and the polynucleotide
XX sequences encoding them. A total of 40 PMMW polypeptides (designated PMMW
XX -1 to PMMW-40) are disclosed. The sequences of the invention are useful
XX for diagnosing a condition or disease associated with the expression of
XX PMMW in a subject, preparing a polyclonal or monoclonal antibody, and
XX generating an expression profile of a sample containing the
XX polynucleotides. The diseases or conditions associated with decreased
XX expression or overexpression of PMMW are cell proliferation disorders
XX (e.g. cancer, atherosclerosis), neurological disorders (e.g. epilepsy,
XX Huntington's disease, stroke), immune/inflammatory disorders, (e.g. AIDS,
XX allergies), developmental disorders (e.g. hypothyroidism, Cushing's
XX syndrome), gastrointestinal or epithelial disorders, and infections. The
XX PMMW polypeptides or their fragments are useful in screening compounds
XX for effectiveness as agonists or antagonists of the polypeptides, or in
XX altering the expression of the target polynucleotide and compounds that
XX specifically bind to, or modulate the activity of the polypeptide.
XX
XX ABU92021-ABU92060 represent the human PMMW polypeptides of the invention
XX
XX Sequence 824 AA;
XX
XX Alignment Scores:
XX Pred. No.: 3e-180 Length: 824
XX Score: 1645.50 Matches: 310
XX Percent Similarity: 86.4% Conservative: 0
XX Best Local Similarity: 86.4% Mismatches: 0
XX Query Match: 76.9% Indels: 49
XX DB: 6 Gaps: 1
XX
XX US-10-825-632-4 (1-1197) x ABU92030 (1-824)
QY 3 TTTGAGGACCAAGACTCCCTTTAGACATCACCCTGCTAGCTAGTACGTACGTAAT 62
Db 466 PheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSerTyrValAsn 485
QY 63 CTGGAGAGGTGACAAAGCTGACTGCGGTGGTACTACATCTTCTGTCATGTCAGT 122
Db 486 ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCysIleSerGln 505
QY 123 CACTGTGACTTCTTTTATAAGTAAAGTATAGTAACCAAGAAATCCACACTGTGTGCCCTT 182
Db 506 HisCysAspPhePheIleSerLysTyrSerAsnGlnLysProHisCysValSerLeu 525
QY 183 TACAAGCTATCAAGTCTCTGAAGATGACCCCAACTTGCACAAACAAAGAAATTTTGGCCACC 242
Db 526 TyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPheTrpAlaThr 545
QY 243 ATTTTGGATTCCAGCAGGTCTCTTCTCTGACTACTACTCTCCAGAAATTTTCTTTTGA 302
Db 546 IleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIlePheSerPheGlu 565
```

QY	303	AGTACTACTGGATTACATTGATGGGATGCTCTACAAGGCTCATGATCTACAGCTTGA	362	FT	Active-site	/note= "Forms part of Ser-Asp-His catalytic triad"
Db	566	SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly	585	FT	Active-site	/note= "Forms part of Ser-Asp-His catalytic triad"
QY	363	AAGAAATATCCCTACTGTGCTTCATATATATGCTGCTCTCAG	404	FN	WO200119866-A1.	/note= "Forms part of Ser-Asp-His catalytic triad"
Db	586	LysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeuValAsnAen	605	XX		
QY	404	-----	404	XX	22-MAR-2001.	
Db	606	ArgPheLysGlyValLysTyrPheArgLeuAenThrLeuAlaSerLeuGlyTyrValVal	625	XX	11-SEP-2000; 2000WO-AU001085.	
QY	404	-----	404	XX	10-SEP-1999; 99AU-00002762.	
Db	626	ValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLys	645	XX	18-FEB-2000; 2000AU-00005709.	
QY	405	-----GGTCAATAGAAATTGACGATCAGGTGGAAGGACTCCCAATATCTAGTCTCT	455	PI	(UNSY) UNIV SYDNEY.	
Db	646	TyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyrLeuAlaSer	665	DR	Abbott CA, Gorell MD;	
QY	456	CGATATGATTTCATTGACTTACATCGTGTGGGCATCCACGGCTGGTCTTATGAGGATAC	515	XX	WPI; 2001-281520/29.	
Db	666	ArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyrGlyGlyTyr	685	PT	N-PSDB; AAC85694.	
QY	516	CTCTCCTCATGCGATTATGACGAGCTCAGATATCTTCAGGTTGCTATTGCTGGGCC	575	PS	New human dipeptidyl aminopeptidase (DPP8) useful for cleaving	
Db	686	LeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAlaGlyAla	705	XX	substrates, identifying inhibitors of DPP8 catalytic activity which have	
QY	576	CGAGTCATCTGTGGATCTTCTATGATACAGGATACACGGACGTTATATGGTCCACCT	635	CC	therapeutic uses, and for detecting activated T cells.	
Db	706	ProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMetGlyHisPro	725	CC	Claim 1; Fig 2; 78pp; English.	
QY	636	GACAGAATGAACAGGGCTATTACTAGGATCTGTGCCATGCAAGAGAAAGTCCCC	695	CC	This sequence represents human dipeptidyl aminopeptidase (DPP8). DPP8 has	
Db	726	AspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGluLysPhePro	745	CC	substrate specificity for H-Ala-Pro-pNA, H-Gly-Pro-pNA and H-Arg-Pro-pNA.	
QY	696	TCTGAACCAAACTGTTACTGCTTACATGTTTCTGGATGAGAAATGTCATTTTGCA	755	CC	Therefore, it is prolyl oligopeptidase and a dipeptidyl peptidase,	
Db	746	SerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAenValHisPheAla	765	CC	because it is capable of hydrolysing the peptide bond C-terminal to Pro	
QY	756	CATACAGTATATTACTAGTGTGTTTGTAGTGAGGGCTGGAAGCCATATGATTTACAGATC	815	CC	in each of these compounds. DPP8 is homologous with human DPPIV. DPP8 is	
Db	766	HisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAspLeuGlnIle	785	CC	useful for cleaving a substrate, and for detecting an activated T cell	
QY	816	TATCCTCAGGAGACACAGCATATAGAGTTCTGTAATCGGAGAACATTATGAACTGCAT	875	CC	which involves measuring the level of DPP8 gene expression in a T cell	
Db	786	TyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyrGluLeuHis	805	CC	The level of DPP8 expression is detected by detecting the amount of DPP8	
QY	876	CTTTGCACTACCTTCAAGAAACCTTGATCAGTATTGCTGCTCTAAAAGTGATA	932	CC	RNA in the cell. It is also useful for identifying a molecule capable of	
Db	806	LeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaLeuLysValIle	824	CC	inhibiting the cleavage of the substrate by DPP8. Molecules identified as	
RESULT 6				CC	inhibiting DPP8 catalytic activity may be useful for treating diarrhoea,	
AAB47187				CC	growth hormone deficiency, lowering glucose levels in non-insulin	
ID	AAB47187	standard; protein; 882 AA.		CC	dependent diabetes mellitus and other disorders involving glucose	
XX	AC	AAB47187;		CC	intolerance, enhancing mucosal regeneration and as immunosuppressants	
XX	DT	29-JUN-2001 (first entry)		XX	Sequence 882 AA;	
XX	DE	Human DPP8.		Alignment Scores:		
XX	XX	Human; dipeptidyl aminopeptidase; DPP8; prolyl oligopeptidase;		Pred. No.:	3.1e-180	Length: 882
KW	KW	dipeptidyl peptidase; DPPIV; T cell; cleavage; diarrhoea;		Score:	1645.50	Matches: 310
KW	KW	growth hormone deficiency; glucose level; mucosal regeneration;		Percent Similarity:	86.4%	Conservative: 0
KW	KW	non-insulin dependent diabetes mellitus; glucose intolerance;		Best Local Similarity:	86.4%	Mismatches: 0
XX	OS	immunosuppression.		Query Match:	76.9%	Indels: 49
XX	XX	Homo sapiens.		DB:	4	Gaps: 1
XX	Key	Location/Qualifiers		US-10-825-632-4 (1-1197) x AAB47187 (1-882)		
PH	Active-site	739		QY	3	TTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTACCTAGTCAGTTAGCTAAAT
FT				Db	524	PheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSerTyrValAsn

```
Db 604 IleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIlePheSerPheGlu 623
Qy 303 AGTACTACTGGATTTTACATTTGATGGGATGCTCTACAGCCTCATGATCTACAGCCTGGA 362
Db 624 SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly 643
Qy 363 AGAATATCTCTACTGCTGCTTTCATATATGCTGCTCTGAG----- 404
Db 644 LysLysTyrProThrValLeuPheIleTyrGlyProGlnValGlnLeuValAsnAsn 663
Qy 404 ----- 404
Db 664 ArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrValVal 683
Qy 404 ----- 404
Db 684 ValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLys 703
Qy 405 -----GGTCAATAGAAATTCACGATCAGGTGGAGGACTCCAAATATCTAGCTTCT 455
Db 704 TyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyrLeuAlaSer 723
Qy 456 CQATATGATTTCATTGACTAGATCGTGTGGGCATCCACGGCTGCTCTATGGAGGATAC 515
Db 724 ArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrSerTyrGlyGlyTyr 743
Qy 516 CTCTCCTCATGGCATTATGACAGAGTTCAGATATCTTCAGGGTGTCTATGTGGGGCC 575
Db 744 LeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAlaGlyAla 763
Qy 576 CAGTCACTCTGTGATCTTCATGATACAGGATACAGGACGTTATATGGTCCACCT 635
Db 764 ProValThrLeuTrpIlePheTyrAspThrThrGluArgTyrMetGlyHisPro 783
Qy 636 GACCAGAAAGACAGGGCTATTACTAGGATCTGTGCCATGCAAGCAGAAAGTCCCC 695
Db 784 AspGlnAsnGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGluLysPhePro 803
Qy 696 TCTGAACCAAACTGTTTACTGCTTACATGTTTCTCGATGAGAAATGTCATTTTGA 755
Db 804 SerGluProAsnArgLeuLeuLeuLeuHisGlyPheLeuAspGluAsnValHisPheAla 823
Qy 756 CATACAGTATATTTACTGATTTTATGAGGGCTGGAAGCCATATGATTTACAGATC 815
Db 824 HisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAspLeuGlnIle 843
Qy 816 TATCCTCAGAGACACAGCATAGACTTCTGAATCGGGAGAACATTATGAATGCAT 875
Db 844 TyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyrGluLeuHis 863
Qy 876 CTTTTGCATCTACCTCAAGAAACCTTGGATCAGTATTGTGCTCTAAAAGTGATA 932
Db 864 LeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaLeuLysValIle 882

RESULT 7
ABG61591
ID ABG61591 standard; protein; 882 AA.
AC ABG61591;
XX
DT 12-AUG-2003 (first entry)
DE Human DPPIV related serine protease DPP-1.
XX
KW Human; serine protease; dipeptidyl peptidase IV-related protein; DPPP;
KW DPPIV; infection; human immunodeficiency virus; HIV-1; HIV-2; pain;
KW diabetes; infertility; obesity; anorexia; Parkinson's disease; stroke;
KW heart failure; hypertension; urinary retention; osteoporosis; cancer;
KW ulcer; allergy; cancer; psychotic disorder; neurological disorder;
KW dyskinesia; reproductive disorder; inflammatory disorder;
KW metabolic disorder.
XX
OS Homo sapiens.
```

```
XX WO200231134-A2.
PN 18-APR-2002.
XX
PD 12-OCT-2001; 2001WO-US031874.
PF
XX 12-OCT-2000; 2000US-0240117P.
PR (FERR ) FERRING BV.
XX
PA Qi S, Akineanya KO, Riviere PJ, Junien J;
PI WPI; 2002-444178/47.
XX N-ESDB; ABK83322.
DR
XX
PT New dipeptidyl peptidase IV-related proteins and nucleic acids encoding
PT the proteins, useful for treating e.g. fungal, bacterial, protozoan and
PT viral infections, cancers, allergies, neurological disorders, or pain.
XX
PS Claim 17; Fig 1; 113pp; English.
XX
CC The present invention relates to the isolation of novel human serine
CC proteases referred to as dipeptidyl peptidase IV (DPPIV)-related proteins
CC (DPRP). The dipeptidyl peptidase IV-related proteins (DPRP) and nucleic
CC acids encoding them are useful for treating infections such as fungal,
CC bacterial, protozoan and viral infections, particularly infections caused
CC by human immunodeficiency virus (HIV-1 or HIV-2), pain, diabetes,
CC precocious puberty, infertility, obesity, anorexia, bulimia, Parkinson's
CC disease, acute heart failure, hypotension, hypertension, urinary
CC retention, osteoporosis, angina pectoris, stroke, ulcers, asthma,
CC allergies, cancers, migraine, vomiting, psychotic and neurological
CC disorders (e.g. anxiety, dementia, or schizophrenia), and dyskinesias.
CC These may also be used in discovering therapeutic agents for the
CC treatment of reproductive, inflammatory and metabolic disorders. ABG61591
CC -ABG61612 represent human DPRP proteins
XX
SQ Sequence 882 AA;
Alignment Scores:
Pred. No.: 3.1e-180 Length: 882
Score: 1645.50 Matches: 310
Percent Similarity: 86.4% Conservatives: 0
Best Local Similarity: 86.4% Mismatches: 0
Query Match: 76.9% Indels: 49
DB: 5 Gaps: 1
US-10-825-632-4 (1-1197) x ABG61591 (1-882)
Qy 3 TTGGAAGGCACCAAGACTCCCTTTAGAGCATACCTGTACGTAGTCAGTTACGTAAT 62
Db 524 PheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSerTyrValAsn 543
Qy 63 CCTGGAGAGGTGACAGGCTGACTGACCGTGTCTACTCACATCTTCTGTCGATCAGTCAG 122
Db 544 ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCysIleSerGln 563
Qy 123 CACTGTGACTTCTTTATAGTAAAGTATAGTAAACCAAGAAATCCACACTGTGTGTCCTT 182
Db 564 HisCysAspPhePheIleSerIleTyrSerAsnGlnLysAsnProHisCysValSerLeu 583
Qy 183 TACAAGCTATCAAGTCTCTGAAGATGACCCCAACTTGCACAAACAAAGAAATTTGGGCCACC 242
Db 584 TyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPheTrpAlaThr 603
Qy 243 ATTTTGGATTCAGCAGGTCTCTCTCTGATATCTCTCCAGAAATTTCTCTTTGAA 302
Db 604 IleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePheSerPheGlu 623
Qy 303 AGTACTACTGGATTTTACATTTGATGGGATGCTCTACAGCCTCATGATCTACAGCCTGGA 362
Db 624 SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly 643
```

```
QY 363 AAGAAATATCCCTACTGCTGCTGTTTCATATATGCTGCTCTCAG-----404
Db 644 LyeLysTyProThrValLeuPheLeuLeuLeuLeuLeuValGlnLeuValAsnAsn 663
QY 404 -----404
Db 664 ArgPheLysGlyValLysTyPheArgLeuAsnThrLeuAlaSerLeuGlyTyrValVal 683
QY 404 -----404
Db 684 valValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLys 703
QY 405 -----GTCAAATAGAAATGACCAATCAGGTGGAAGACTCCCAATATCTAGCTTCT 455
Db 704 TyrLysMetGlyGlnIleGluLeuAspGlnValGluGlyLeuGlnTyrLeuAlaSer 723
QY 456 CGATATGATTTCATTGACTTAGATCGTGTGGGCATCCACGGCTGGTCTATGGAGATAC 515
Db 724 ArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrSerTyrGlyGlyTyr 743
QY 516 CTCTCCCTGATGGCATTAAATGCAGAGGTTCAGATATCTTCAGGGTTGCTATTGTGGGCC 575
Db 744 LeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAlaGlyAla 763
QY 576 CCACTCACCTCTGTGATCTTCTATGATACAGATACACGGAACGTTATATGGTCCACT 635
Db 764 ProValThrLeuTyrIlePheTyrAspThrGlyTyrThrGluArgTyrMetGlyHisPro 783
QY 636 GACCAAGTACAGGGCTATTACTAGATCTGTGCCATCGACGAGAAAGTTCCCTCC 695
Db 784 ArgGlnAsnGluGlnGlyTyrLeuGlySerValAlaMetGlnAlaGluLysPhePro 803
QY 696 TCTGAACCAAAATCGTTTACTGCTTACATCGTTTCTGAGTGAAGATGCCATTTGCA 755
Db 804 SerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluLeuValHisPheAla 823
QY 756 CATACACGATATTACTAGTATTTTGTAGTGGCTCGAAAGCCATATGATTTACAGATC 815
Db 824 HisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAspLeuGlnIle 843
QY 816 TATCTCTAGGAGACACAGCATAGAGTTCTCGAATCGGGAGAACATTTATGAATGCAT 875
Db 844 TyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyrGluLeuHis 863
QY 876 CTTTGTGCACTACTTCAAGAAACCTTGGATCAGTATTGCTGCTCTAAAGTGATA 932
Db 864 LeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeuLysValIle 882

RESULT 8
AAG78415
ID AAG78415 standard; protein; 882 AA.
XX
AC AAG78415;
XX
DT 12-APR-2002 (first entry)
XX
DE Amino acid sequence of 21953 human prollyl oligopeptidase.
XX
KW 21953 prollyl oligopeptidase; antibody; proline; endopeptidase; cancer;
KW cardiovascular disease; autoimmune disease; atopic allergy;
KW neuronal disorder; vascular disorder; prostate disorder; cytostatic;
KW antidiabetic; antiarthritic; antiasthmatic; antiinflammatory;
KW diabetes mellitus; arthritis; multiple sclerosis; asthma;
KW Grave's disease; neuronal disorder; demyelinating disease.
XX
OS Homo sapiens.
XX
PN WO2001/9473-A2.
XX
PD 25-OCT-2001.
XX
PF 11-APR-2001; 2001WO-US040483.
XX
```

```
PR 18-APR-2000; 2000US-0197508P.
XX (MILL-) MILLENNIUM PHARM INC.
XX
PI Meyers RA, Williamson M;
XX
DR WPI; 2002-034353/04.
DR N-ESDB; AAH99934.
XX
PT New polypeptides 21953, member of human prollyl oligopeptidase family,
PT useful as diagnostic targets and therapeutic agents for controlling
PT cancer, lymphoma and leukemia.
PS
XX Claim 1; Page 102-103; 121pp; English.
XX
CC This invention relates to an isolated 21953 human prollyl oligopeptidase.
CC Which is cytosolic, antidiabetic, antiarthritic, neuroprotective,
CC antihypertensive, dermatological, antiparasitic, antiasthmatic,
CC ophthalmological, antiinflammatory, nootropic, antiparkinsonian,
CC anticonvulsant, gynaecological vasotrophic, antianxiety, cardiac,
CC antithrombotic, anorectic and metabolic in its action. Uses include
CC gene therapy, expression or activity of 21953 protein modulator, it is
CC useful for identifying a compound which binds to it and can be used in
CC preventing, treating or detecting a cellular proliferative or
CC differentiative disorder. The 21953 molecules can act as novel diagnostic
CC targets and therapeutic agents for controlling disorders associated with
CC the aberrant activity or degradation of peptide hormones e.g., disorders
CC associated with cell differentiation and proliferation such as cancer,
CC immune function, reproductive, neurological and cardiovascular function.
CC The 21953 molecules are thus useful for treating and preventing cellular
CC proliferative and differentiative disorders, haematopoietic neoplastic
CC disorders, immune disorders such as autoimmune diseases, diabetes
CC mellitus, arthritis, multiple sclerosis, asthma, Grave's disease,
CC neuronal disorders, demyelinating diseases, vascular disorders and
CC metabolism or pain disorders. This sequence represents the amino acid
CC sequence of 21953 human prollyl oligopeptidase
XX
SQ Sequence 882 AA;

Alignment Scores:
Pred. No.: 3.1e-180 Length: 882
Score: 1645.50 Matches: 310
Percent Similarity: 86.4% Conservative: 0
Best Local Similarity: 86.4% Mismatches: 0
Query Match: 76.9% Indels: 49
DB: 5 Gaps: 1

US-10-825-632-4 (1-1197) x AAG78415 (1-882)
QY 3 TTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGACGTAGTCAGTTACGTAAT 62
Db 524 PheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSerTyrValAsn 543
QY 63 CTTGGAGAGGTGACAGGCTGACTGCGTGGCTACTCACATTCTTGTGCTCAGTCAG 122
Db 544 ProGluGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysIleSerGln 563
QY 123 CACTGTGACTCTCTTTATAGTAAAGTATAGTAAACGAGAAATCCACACTGTGTCTCCCTT 182
Db 564 HisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeu 583
QY 183 TACAAGCTATCAAGTCTCTGAAGATCACCACTTGTCAAAACAAAGGAATTTTGGCCACC 242
Db 584 TyrLysLeuSerSerProGluAspProThrCysLysThrLysGluPheTrpAlaThr 603
QY 243 ATTTTGGATTACAGAGGCTCTCTCTCTGACTATATCTCTCCAGAAATTTTCTTTTGA 302
Db 604 IleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIlePheSerPheGlu 623
QY 303 AGTACTACTGGATTTTACATTTGATGGGATGCTCTACAGACCTCATGATCTACAGCCTGGA 362
Db 624 SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly 643
```

```
QY 363 AAGAAATATCTACTGCTGCTGTTTCATATATATGTTGCTCTCAG----- 404
Db 644 LysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeuValAsnAsn 663
QY 404 ----- 404
Db 664 ArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrValVal 693
QY 404 ----- 404
Db 684 ValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLys 703
QY 405 ----- 404
Db 704 TyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyrLeuAlaSer 723
QY 456 CGATATGATTTCATTGACATTAGCTGTGGGCGATCCACGGCTGCTCTATGGAGGATAC 515
Db 724 ArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyIleGlyTyrSerTyrGlyTyr 743
QY 516 CTCTCCCTGATGGCATTATATGAGAGGTTCAGATATCTTCAGGGTTGCTATTGCTGGGGCC 575
Db 744 LeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAlaGlyAla 763
QY 576 CAGTCACTCTGTGGATCTTCATGATACAGATACAGGAAAGCTTATATGGGTCAACCT 635
Db 764 ProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMetGlyHisPro 783
QY 636 GACCAGAATGACAGGGCTATTACTTAGGATCTGTGGCCATGCGCAAGCAGAGAAAGTCCCC 695
Db 784 ArgGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGluLysPhePro 803
QY 696 TCTGAACCAAACTGTTACTGCTTACATGTTTCTTGGATGAGAAATGTCATTTTGA 755
Db 804 SerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnValHisPheAla 823
QY 756 CATACAGTATATTACTGATGTTTTTGTAGGGCTGGAAAGCCATATGATTTACAGATC 815
Db 824 HisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAspLeuGlnIle 843
QY 816 TATCCTCAGGAGACAGCATAGAGTTCCTGAATCGGGAGAACATTTATGAATGCAT 875
Db 844 TyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyrGluLeuHis 863
QY 876 CTTTTCACATCTCTCAAGAAACCTTGGATCAGTATTGCTGCTCTAAAAGTGATA 932
Db 864 LeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaLeuLysValIle 882

RESULT 9
AAE24170
ID AAE24170 standard; protein; 882 AA.
XX
AC AAE24170;
XX
XX 23-SEP-2002 (first entry)
XX
XX Human dipeptidyl peptidase 8 (DPP8) protein.
XX
XX Human; dipeptidyl peptidase; DPP; neoplasia; type II diabetes; cirrhosis;
XX autoimmunity; human immuno deficiency virus; HIV infection; cytostatic;
XX graft rejection; antidiabetic; antiinflammatory; immunosuppressive;
XX antiviral; enzyme.
XX
XX Homo sapiens.
XX
XX WO200234900-A1.
XX
XX 02-MAY-2002.
XX
XX 29-OCT-2001; 2001WO-AU001388.
XX
XX 27-OCT-2000; 2000AU-00001078.
XX
XX
```

```
PA (UNSY ) UNIV SYDNEY.
XX
XX PI Abbott CA, Gorrell MD;
XX
XX DR WPI; 2002-454646/48.
XX DR N-PSDB; RAD38956.
XX
XX New dipeptidyl peptidase (DPP) peptides, useful for screening inhibitors
XX of DPP catalytic activity, which may be employed to treat e.g. neoplasia,
XX type II diabetes, cirrhosis, autoimmunity, graft rejection and HIV
XX infection.
XX
XX Example; Fig 1; 91pp; English.
XX
XX The present invention relates to dipeptidyl peptidase (DPP) proteins and
XX polynucleotides encoding such proteins. The DPP peptides are useful for
XX screening inhibitors of ppp catalytic activity. The inhibitors are useful
XX for treating neoplasia, type II diabetes, cirrhosis, autoimmunity, graft
XX rejection and HIV (human immuno deficiency virus) infection. The present
XX sequence is human DPP8 protein
XX
XX Sequence 882 AA;
XX
XX Alignment Scores:
XX Pred. No.: 3.1e-180 Length: 882
XX Score: 1645.50 Matches: 310
XX Percent Similarity: 86.4% Conservative: 0
XX Best Local Similarity: 86.4% Mismatches: 0
XX Query Match: 76.9% Indels: 49
XX DB: 5 Gaps: 1
XX
XX US-10-825-632-4 (1-1197) x AAE24170 (1-882)
XX
QY 3 TTTGAAGGACCAAGACTCCCTTTAGAGCATCACCTGTAGTACGTAGTCAGTTACGTAAT 62
Db 524 PheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSerTyrValAsn 543
QY 63 CCTGGAGAGTGCAAGGCTGACTGACCGTGGCTACTCACATTCCTTGTGTCATGTCAG 122
Db 544 ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCysIleSerGln 563
QY 123 CACTGTGACTCTTTTATATAGTATAGTATACACGAGAGATCCACACTGTGTGCTCCCTT 182
Db 564 HisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeu 583
QY 183 TACAAGCTATCAAGTCTCTGAAGATGACCACTGCAAAACAAAGAAATTTGGSCCACC 242
Db 584 TyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPheTrpAlaThr 603
QY 243 ATTTTGGATTACAGAGGCTCTCTTCTGACTATATCTCTCCAGAAATTTTCTTTTGAA 302
Db 604 IleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIlePheSerPheGlu 623
QY 303 AGTACTACTGGATTTACATTTGATGGGATGCTCTACAGCCCTCATGATCTACAGCCTGGA 362
Db 624 SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly 643
QY 363 AAGAAATATCTACTGCTGCTTTTCATATATGTTGTTGCTCTCAG----- 404
Db 644 LysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeuValAsnAsn 663
QY 404 ----- 404
Db 664 ArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrValVal 683
QY 404 ----- 404
Db 684 ValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLys 703
QY 405 -----GGTCAATAGAAATTGACGATCAGGTGGAGGACTCCCAATATCTAGCTTCT 455
Db 704 TyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyrLeuAlaSer 723
```


CC antiasthmatic, nephrotropic, antiarthritic, hepatotropic,
CC neuroprotective, nootropic, antibacterial, virucide, antiparasitic,
CC relaxant and anticonvulsant. In addition, they are useful in screening
assays to identify small molecules that modulate or inhibit, for example,
CC neurogenesis, wound healing and angiogenesis. The nucleic acids are also
used as in chromosome mapping, tissue typing, preventive medicine and
CC pharmacogenomics. This polypeptide is a homologue of a human NOVX protein
CC of the invention.
XX
SQ Sequence 882 AA;

Alignment Scores:
Pred. No.: 3 1e-180 Length: 882
Score: 1645.50 Matches: 310
Percent Similarity: 86.4% Conservativity: 0
Best Local Similarity: 86.4% Mismatches: 0
Query Match: 76.9% Indels: 49
DB: 5 Gaps: 1

US-10-825-632-4 (1-1197) x ADI117086 (1-882)

QY 3 TTGAAGGACCAAGAGCTCCCTTAGAGCATCACCTGTACGTACTACGTAAAT 62
DB 524 PheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSerTyrValAsn 543
QY 63 CTGAGAGGTCAAGGCTGACTGACCGTGGCTACTCACATCTTCTGCTGCATCAG 122
DB 544 ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCysLeuSerGln 563
QY 123 CACTGTGACTCTTTTATAGTAAGTATAGTAAACAGAGAATCCACATGTGTGCTCCCTT 182
DB 564 HisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeu 583
QY 183 TACAAGCTATCAAGTCTGGAAGATGACCCAACTTGCMAAAACAAAGAAATTTGGCCACC 242
DB 584 TyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPheTrpAlaThr 603
QY 243 ATTTTGGATTACAGAGGCTCTCTCTGACTATATCTCTCCAGAAATTTCTCTTTGAA 302
DB 604 IleLeuAspSerAlaGlyProLeuProAspTyrThrProGluLysPheSerPheGlu 623
QY 303 AGTACTACTGGATTTTACATTTGATGGGATGCTCTACAGGCTCATGATCTACAGCTGGA 362
DB 624 SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly 643
QY 363 AAGAAATATCTACTGTGCTGTTCATATATGCTGTGCTCTCAG----- 404
DB 644 LysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeuValAsnAsn 663
QY 404 ----- 404
DB 664 ArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrValVal 683
QY 404 ----- 404
DB 684 ValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLys 703
QY 405 -----GGTCAATAGAAATTGACATCAGTGGAGAGCTCCGAATATCTAGCTTCT 455
DB 704 TyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyrLeuAlaSer 723
QY 456 CGATATGATTTTCACTTACATTCGTGTGGCATCCACGGCTGGTCTCTATGAGAGATAC 515
DB 724 ArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrSerTyrGlyGlyTyr 743
QY 516 CTCTCCCTGATGGCATTAATCAGAGGTCAGATATCTTTCAGGGTGTCTATTGCTGGGGCC 575
DB 744 LeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAlaGlyAla 763
QY 576 CCAGTCACTCTGTGGATCTTCTATGATACAGGATACAGGAACCTTATATGGGTCAACCT 635
DB 764 ProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMetGlyHisPro 783

QY 636 GACCAGAAATGAACAGGGCTATTACTTACGATCTCTGGCCATGCAACAGAAAGTTCCCC 695
DB 784 AspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGluLysPhePro 803
QY 696 TCTGAACCAAAATCGTTTACTCTCTTACATGTTTCTGATGAGAAATGTCATTTTGA 755
DB 804 SerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnValHisPheAla 823
QY 756 CATACCAGTATATTACTGAGTGTCTTGTAGTGGCTCGAAAGCCATATGATTTACAGATC 815
DB 824 HisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAspLeuGlnIle 843
QY 816 TATCTCTCAGGAGAGACACAGCATTAAGATTCCTCAATCGGAGAACATTAAGACTGCAT 875
DB 844 TyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHistyrGluLeuHis 863
QY 876 CTTTTCGACTACCTTCAAGAAACCTTGGATCAGCTATTGCTCTTAAAGATGATA 932
DB 864 LeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaLeuLysValIle 882

RESULT 12

ABU07720
ID ABU07720 standard; protein; 882 AA.

XX AC ABU07720;

XX DT 19-MAY-2003 (first entry)

XX DE Human serine protease HIPHUM46.

XX KW Human; enzyme; HIPHUM46; serine protease; gene therapy; osteoarthritis;
KW serine protease activity modulation; dipeptidyl peptidase activity;
KW musculoskeletal disease; Hepatitis B virus infection; myotonic dystrophy;
KW Alzheimer's disease; parasupranuclear palsy; Huntington's disease;
KW amyotrophic lateral sclerosis; malabsorption syndrome; lung disease;
KW irritable bowel syndrome; type I diabetes; fascial incontinence; anaemia;
KW haemorrhoid; proctitis; rectal polyp; small bowel tumour; dyslexia;
KW colorectal tumour; ceroid lipofuscinosis; allergic encephalomyelitis;
KW multiple sclerosis.

XX KW Homo sapiens.

XX OS Key Location/Qualifiers
XX Region 259..260
XX /note= "Paired glutamates of the beta propeller domain"

FT Active-site 739 /label= Catalytic_serine_residue

FT Active-site 817 /label= Catalytic_aspartate_residue

FT Active-site 849 /label= Catalytic_histidine_residue

XX GB2374869-A.

XX 30-OCT-2002.

XX 22-JAN-2002; 2002GB-00001404.

XX 23-JAN-2001; 2001GB-00001760.

XX (GLAX) GLAXO GROUP LTD.

XX Edbrooke MR, Lewis AP;

XX WPI; 2003-150703/15.

XX N-PSDB; ABX12255.

XX Identifying modulators of serine protease activity useful for treating
XX musculoskeletal diseases, by contacting cell expressing a novel serine
XX protease polypeptide with a compound and monitoring serine protease
XX activity.

PS Claim 10; Page 26-29; 38pp; English.

XX The invention relates to a method of identifying a substance that
CC modulates serine protease activity, comprising contacting a cell such as
CC a neuronal cell, lung cell, intestinal cell or a cell infected with a
CC virus, expressing a serine protease polypeptide (HIPHUM 46), or its
CC variant having dipeptidyl peptidase activity, or a serine protease
CC isolated from the cell with a test substance, and monitoring for serine
CC protease activity. The method is useful for identifying a substance that
CC modulates serine protease activity. A modulator of the serine protease is
CC useful in the manufacture of a medicament for treatment or prophylaxis of
CC a musculoskeletal disease e.g. osteoarthritis, Hepatitis B virus
CC infection, Alzheimer's disease, paraspranuclear palsy, myotonic
CC dystrophy, Huntington's disease or amyotrophic lateral sclerosis.
CC Additional disease that may be treated using modulators of the serine
CC protease include malabsorption syndromes, irritable bowel syndrome, lung
CC disease, type I diabetes, faecal incontinence, haemorrhoids, proctitis,
CC rectal polyps, small bowel tumours, colorectal tumours, anaemia,
CC dyslexia, ceroid lipofuscinosis, allergic encephalomyelitis, and multiple
CC sclerosis. The present sequence represents the amino acid sequence of the
CC human serine protease HIPHUM46
XX
SQ Sequence 882 AA;

Alignment Scores:
Pred. No.: 3.1e-180 Length: 882
Score: 1645.50 Matches: 310
Percent Similarity: 86.4% Conservative: 0
Best Local Similarity: 86.4% Mismatches: 0
Query Match: 76.9% Indels: 49
DB: 6 Gaps: 1

US-10-825-632-4 (1-1197) x ABU07720 (1-882)

QY 3 TTGAGGACCAAGACTCCCTTTAGACATCACCTGTACGTAGTCAGTACGTAAAT 62
DB 524 PHEGLUGLYTHRLYSANSPERPROLEUGLHSHISLEUTYRVALSERTYRVALAM 543
QY 63 CTGGAGAGGTGACAGGCTGACTGACCGTGGCTACTCACATCTTGTCTGCATCAGTCAG 122
DB 544 PROGLYGLUVALTHRARGLEUTHRASPARGGLYTYRSERHISSERCYSSILLESERGLN 563
QY 123 CACTGTGACTCTTTATAGTATAGTATAGTATAGTATAGTATAGTATAGTATAGTATAGT 182
DB 564 HISCYASPPHEPHEILESERLYSTYRSERASNGLNLYSASNPROHISCYSSVALSERLEU 593
QY 183 TACAACCTATCAAGTCTGACATGACCCCACTTGCAAAACAAAGGAATTTTGGCCACC 242
DB 584 TYRLYSLEUSERSERPROGLUASPPROTHRCYSLYSTHRLYSGLUPHETRPALATHR 603
QY 243 ATTTTGGATTCAGCAGGTCCTCTCTGACTATATCTCTCCAGAAATTTTCTCTTTGAA 302
DB 604 ILELEUASPSERALAGLYPROLEUPROASPTYRTHRPROGLIULIIPHESERPHEGLU 623
QY 303 AGTACTACTGATTTACATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 362
DB 624 SERTHRTHRGLYPHETHRLEUTYRGLYMETLEUTYRYSPROHISASPLEUGLINPROGLY 643
QY 363 AAGAAATATCTACTGTCTGTCTATATATATATATATATATATATATATATATATATATAT 404
DB 644 LYSLEYTYRPROTHRVALHEUPHEILETYRGLYGLYPROGLINVALGLINLEUVALASN 663
QY 404 ----- 404
DB 664 ARGPHELYSGLYVALYSTYRPHEARGLAUSNTHRLAUSERLEUGLYTYRVALVAL 683
QY 404 ----- 404
DB 684 VALVALILEASPNARGGLYSERCYSHISARGGLYLEULYSPHEGLUGLYALAPHELYS 703
QY 405 -----GGTCAATAGAAATTCAGATCAGGTGGAGGACTCCAAATATCTAGTTCT 455
DB 704 TYRLYSMETGLYGLINILEGLUILEASPSGLINVALGLUGLYLEUGINTYRLEUVALASER 723

QY 456 CGATATGATTTTCTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTT 515
DB 724 ARGTYRASPHEILEASPLEUASPARGVALGLYVILIEHISGLYTRPSERTYRGLYGLYTYR 743
QY 516 CTCTCCCTGATGCATTAATAGCAGAGGTGAGATATCTTCTCAGGGTTGCTTATTCCTGGGCC 575
DB 744 LEUSERLEUMETALALEUUMETGLNARGSERASPLEIPHEARGVALALALEALAGLYALA 763
QY 576 CCAGTCACTCTGTGGATCTTCTATATGATACAGGATACAGGAACGTTATATGGTCCACCT 635
DB 764 PROVALTHRLLEUTRPILEPHETYRASPETHRGLYTYRTHRGUARGTYRMETGLYHISPRO 783
QY 636 GACCAGATGAACAGGGCTATTACTTACGATCTGTGGCCATCGACAGCAGAAAGTTCCCC 695
DB 784 ASPGLNASNGLUGLNGLYTYRTHRGUARGSERVALALAMETGLNHALAGLULYSPEPRO 803
QY 696 TCTGAACCAATCTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACT 755
DB 804 SERGLUPROASNARGLEULEULEULEULEULEULEULEULEULEULEULEULEULEULE 823
QY 756 CATACCAAGTATATTTACTGAGTTTTTTAGTGAGGGCTGGAAAGCCATATGATTTACAGATC 815
DB 824 HISTHRSERILELEULEULEULEULEULEULEULEULEULEULEULEULEULEULEULE 843
QY 816 TATCTCTCAGCAGACACACAGCATATAGAGTTCTCTGATCGGAGAACATTATGACTGCAT 875
DB 844 TYRPROGLINGLUARGHISERILEARGVALPROGLIUSERGLYGLYHISGLYGLYHIS 863
QY 876 CTTTGTGACTTACTTCAAGAAACCTTGGATCACGTATTGCTGTCTCTAAAGTGTATA 932
DB 864 LEULEUHISTYRLEUGLNGLEULEULEULEULEULEULEULEULEULEULEULEULEULE 882
RESULT 13
ADIL17085
ID ADIL17085 standard; protein; 883 AA.
XX
AC ADIL17085;
DT
DT 15-APR-2004 (first entry)
XX
DE Murine NOVX protein homologue SeqID 621.
XX
KW mouse; murine; NOVX; cardiomyopathy; atherosclerosis; cancer; diabetes;
KW inflammation; autoimmune disorder; allergy; blood disorder;
KW acquired immunodeficiency syndrome; AIDS; obesity; asthma;
KW immunoglobulin (Ig)A nephropathy; cirrhosis; arthritis;
KW Alzheimer's disease; infect.
XX
OS Mus musculus.
XX
PN WO200268649-A2.
XX
PD 06-SEP-2002.
XX
PF 31-JAN-2002; 2002WO-US002785.
XX
PR 31-JAN-2001; 2001US-0265395P.
PR 31-JAN-2001; 2001US-0265412P.
PR 31-JAN-2001; 2001US-0265514P.
PR 31-JAN-2001; 2001US-0265517P.
PR 02-FEB-2001; 2001US-0266406P.
PR 05-FEB-2001; 2001US-0266767P.
PR 07-FEB-2001; 2001US-0266975P.
PR 07-FEB-2001; 2001US-0267057P.
PR 08-FEB-2001; 2001US-0267459P.
PR 09-FEB-2001; 2001US-0267823P.
PR 15-FEB-2001; 2001US-0268974P.
PR 26-FEB-2001; 2001US-0271664P.
PR 27-FEB-2001; 2001US-0271839P.
PR 27-FEB-2001; 2001US-0271855P.
PR 02-MAR-2001; 2001US-0272788P.
PR 12-MAR-2001; 2001US-0273046P.
PR 14-MAR-2001; 2001US-0275925P.

PR 14-MAR-2001; 2001US-0275947P.
PR 14-MAR-2001; 2001US-0275950P.
PR 14-MAR-2001; 2001US-0275989P.
PR 15-MAR-2001; 2001US-0276448P.
PR 15-MAR-2001; 2001US-0276450P.
PR 16-MAR-2001; 2001US-0276397P.
PR 16-MAR-2001; 2001US-0276397P.
PR 20-MAR-2001; 2001US-0278652P.
PR 20-MAR-2001; 2001US-0278652P.
PR 26-MAR-2001; 2001US-0278775P.
PR 26-MAR-2001; 2001US-0278778P.
PR 29-MAR-2001; 2001US-0279882P.
PR 29-MAR-2001; 2001US-0279884P.
PR 30-MAR-2001; 2001US-0280147P.
PR 11-APR-2001; 2001US-0282992P.
PR 11-APR-2001; 2001US-0283083P.
PR 20-APR-2001; 2001US-0285133P.
PR 23-APR-2001; 2001US-0285749P.
PR 03-MAY-2001; 2001US-0288327P.
PR 03-MAY-2001; 2001US-0288504P.
PR 29-MAY-2001; 2001US-0294047P.
PR 30-MAY-2001; 2001US-0294473P.
PR 08-JUN-2001; 2001US-0296964P.
PR 18-JUN-2001; 2001US-0298959P.
PR 19-JUN-2001; 2001US-0299324P.
PR 13-AUG-2001; 2001US-0312020P.
PR 16-AUG-2001; 2001US-0312889P.
PR 16-AUG-2001; 2001US-0312908P.
PR 21-AUG-2001; 2001US-0313390P.
PR 28-AUG-2001; 2001US-0313470P.
PR 31-AUG-2001; 2001US-0316447P.
PR 07-SEP-2001; 2001US-0318115P.
PR 07-SEP-2001; 2001US-0318118P.
PR 12-SEP-2001; 2001US-0318740P.
PR 18-SEP-2001; 2001US-0323579P.
PR 18-OCT-2001; 2001US-0330245P.
PR 18-OCT-2001; 2001US-0330308P.
PR 14-NOV-2001; 2001US-0332701P.
XX PA (CURA-) CURAGEN CORP.
XX
XX Tchernev VT, Spytek KA, Zerhusen BD, Patturajan M, Shinkets RA;
PI Li L, Gangolli EA, Padigaru M, Anderson DW, Rastelli L, Miller CE;
PI Gerlach VL, Taupier RJ, Gusev VV, Colman SD, Wolenc AR, Pena CE;
PI Furtak K, Grosse WM, Alsobrook JP, Lepley DM, Rieger DK, Burgess CE;
XX WPI; 2002-706998/76.
XX
XX New NOVX polypeptides and nucleic acids, useful for preventing or
PT treating NOVX-associated disorders, e.g. cancer, cardiomyopathy,
PT atherosclerosis, or diabetes, and in chromosome mapping, tissue typing or
PT pharmacogenomics.
XX
XX Disclosure; SEQ ID NO 621; 1498pp; English.
XX
XX This invention relates to a novel nucleic acids, and encoded polypeptides
CC thereof, which have properties related to the stimulation of biochemical
CC or physiological responses in a cell, tissue, organ or organism.
CC Specifically, it refers to the use of biologically active fragments for
CC diagnostic and prognostic assays and furthermore in the treatment of
CC diverse pathological conditions. The present invention describes novel
CC human and murine NOVX proteins, as well as methods to modulate their
CC expression using antisense oligos, ribozymes and peptide nucleic acids.
CC The NOVX polypeptides, polynucleotides and antibodies are useful in
CC treating or preventing NOVX-associated disorders, e.g. cardiomyopathy,
CC atherosclerosis, cancer and diabetes. Furthermore, they may be used in
CC treating or preventing diseases such as inflammation, autoimmune
CC disorders, allergies, blood disorders, acquired immunodeficiency syndrome
CC (AIDS), obesity, asthma, immunoglobulin (Ig) A nephropathy, cirrhosis,
CC arthritis, Alzheimer's disease, infections, stroke, muscular dystrophy
CC and epilepsy. Accordingly, these molecules have many activities including
CC cytostatic, cardiant, antiinflammatory, immunosuppressive, antiallergic,
CC haemostatic, anti-HIV, antidiabetic, antiarteriosclerotic, anorectic,
CC antiasthmatic, nephrotropic, antiarthritic, hepatotropic, 695
CC
CC neuroprotective, nootropic, antibacterial, virucide, antiparasitic,
CC relaxant and anticonvulsant. In addition, they are useful in screening
CC assays to identify small molecules that modulate or inhibit, for example,
CC neurogenesis, wound healing and angiogenesis. The nucleic acids are also
CC used as in chromosome mapping, tissue typing, preventive medicine and
CC pharmacogenomics. This polypeptide is a homologue of a human NOVX protein
CC of the invention.
XX
XX SQ Sequence 883 AA;
XX
XX Alignment Scores:
XX Pred. No.: 1,31e-175 Length: 883
XX Score: 1605.50 Matches: 302
XX Percent Similarity: 85.0% Conservative: 3
XX Best Local Similarity: 84.1% Mismatches: 5
XX Query Match: 75.0% Indels: 49
XX DB: 5 Gaps: 1
XX
XX US-10-825-632-4 (1-1197) x ADI17085 (1-883)
QY 3 TTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTAGTACGTTCAGTTACGTAAAT 62
DB 525 PheGluGlyThrIysAspSerProLeuGluHisIleuTyValThrSerTyAlaAsn 544
QY 63 CCTGGAGAGGTGACAGAGGCTGACTGACCGTGGCTACTCACATTCTTCTGCTGACAGTCAG 122
DB 545 ProGlyGluValValArgLeuThrAspArgGlyTySerHisSerCysCysLeuSerArg 564
QY 123 CACTGTGACTTCTTTATAGTAAGTATAGTAACCGAGAGAAATCCACACTGTGTCTTCCTT 182
DB 565 HisCysAspPhePheIleSerLysTySerAenGlnLysAsnProHisCysValSerLeu 584
QY 183 TACAAGCTATCAAGTCTCTGAAGATGACCCAACTTGCAAAAACAAAGGAATTTTGGGCGACC 242
DB 585 TyLysLeuSerSerProGluAspProValHisLysThrLysGluPheTrpAlaThr 604
QY 243 ATTTTGGATTGACAGAGTCTCTCTCTGACTATATCTCTCCAGAAATTTCTCTTTTCAA 302
DB 605 IleLeuAspSerAlaGlyProLeuProAspTyThrProProGluIlePheSerPheGlu 624
QY 303 AGTACTACTGGATTTTACATTTGATGGGATGCTTACAGAGCCTCATGATCTACAGCCTGGA 362
DB 625 SerThrThrGlyPheThrLeuTyGlyMetLeuTyLysProHisAspLeuGlnProGly 644
QY 363 AAGAAATATCTACTGTCTGTTCATATATGTTGCTGCTCTCAG----- 404
DB 645 LysLysTyProThrValLeuPheIleTyGlyGlyProGlnValGlnLeuValAsnAsn 664
QY 404 ----- 404
DB 665 ArgPheLysGlyValLysTyPheArgLeuAsnThrLeuAlaSerLeuGlyTyValVal 684
QY 404 ----- 404
DB 685 ValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLys 704
QY 405 -----GGTCAATAGAAATTCACATCAGTGGAGGACTCCGACTATCTAGCTTCT 455
DB 705 TyLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyLeuAlaSer 724
QY 456 CGATATGATTTCACTTACATTCGTGTGGGCATCCACGGCTGCTGCTATGAGAGATAC 515
DB 725 GlnTyAspPheIleAspLeuAspArgValGlyIleHisGlyTyPheSerTyGlyTy 744
QY 516 CTCTCCCTGATGGCAATTAATCAGAGGTCAGATATCTTCAGGGTTGCTATTGCTGGGGCC 575
DB 745 LeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAlaGlyAla 764
QY 576 CCAGTCACCTCTGTGGATCTTCTATGATACAGGATACACGGAACTTATATGGGTTCACCT 635
DB 765 ProValThrLeuTrpIlePheTyAspThrGlyTyThrGluArgTyMetGlyHisPro 784
QY 636 GACCAGAATGAACAGGGCTATTACTTAGGATCTGTGGCCATCTGTGCAAGCAAGAAAGTTCCCC 695

Db 785 AppGlnAsnGluGlnGlyTyrLeuGlySerValAlaMetGlnAlaGluLysPhePro 804
Qy 696 TCTGAACCAATCGTTTACTGCTTCTTACATGCTTCTCGGAGAGATGTCATTTTGCA 755
Db 805 SerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnValHisPheAla 824
Qy 756 CATACAGTATATTACTGAGTGTGTTTATGAGGGCTGGAAGCCATATGATTACAGATC 815
Db 825 HisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAspLeuGlnIle 844
Qy 816 TATCCTCAGAGAGACACAGCATAGAGTCTCTGAATCGGAGAACATATTGACATGCAT 875
Db 845 TyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyrGluLeuHis 864
Qy 876 CTTTTGCATCTACCTCAAGAAAACCTTGATCAGCATGCTGCTCTCTAAAAGTGATA 932
Db 865 LeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaLeuLeuValIle 883
RESULT 14
ADE78977
ID ADE78977 standard; protein; 493 AA.
XX
AC ADE78977;
DT 29-JAN-2004 (first entry)
XX
DE Human protein modification and maintenance molecule (PMMW) -15.
XX
KW protein modification and maintenance molecule; PMMW;
KW protein modification; protein maintenance; protein function;
KW protein conformation; protein stabilisation; protein degradation; kinase;
KW phosphatase; protease; protease inhibitor; isomerase; transferase;
KW molecular chaperone; anti-HIV; anti-allergic; antiinflammatory;
KW antianaemic; antiparkinsonian; nootropic; anticonvulsant;
KW antiarteriosclerotic; antiasthmatic; immunosuppressive; antithyroid;
KW cytostatic; hepatotropic; dermatological; antidiabetic; nephrotropic;
KW antigout; thymimetic; neuroprotective; osteopathic; antiarthritic;
KW antiparasitic; antihelminthic; antipsoriatic; uropathic; ophthalmological;
KW antirheumatic; haemostatic; antibacterial; virucide; protozoacide;
KW fungicide; gene therapy; cell proliferative disorder; arteriosclerosis;
KW hepatitis; polycythaemia vera; psoriasis; primary thrombocytopaenia;
KW cancer; developmental disorder; anaemia; mental retardation;
KW neurological disorder; Alzheimer's disease; Parkinson's disease;
KW epilepsy; autoimmune disorder; inflammatory disorder; AIDS; allergies;
KW asthma; autoimmune thyroiditis; Crohn's disease; diabetes mellitus;
KW glomerulonephritis; Goodpasture's syndrome; multiple sclerosis;
KW arthritis; osteoporosis; pancreatitis; Sjogren's syndrome;
KW microbial infection; human.
XX
OS Homo sapiens.
XX
PN WO2003063688-A2.
XX
PD 07-AUG-2003.
XX
PF 23-JAN-2003; 2003WO-US0002500.
XX
PR 25-JAN-2002; 2002US-0351928P.
PR 23-FEB-2002; 2002US-0359903P.
PR 21-MAR-2002; 2002US-0366837P.
XX
XX (INCYTE GENOMICS INC.
XX Hafalia AJA, Li JX, Gorvad AE, Chawla NK, Sprague WW, Lee SY,
PI Chang H, Elliott VS, Ramkumar J, Khare R, Emerling EM, Kable AE,
PI Tang YT, Yue H, Gietzen KO, Lee S, Swarnakar A, Baughn MR,
PI Wilson AD, Jin P, Chien D, Hawkins PR, Jiang X, Jackson AA,
PI Bhatia U, Burrill JD, Blake JJ, Ho A, Zheng W, Ison CH, Marquis JP,
PI Tran UK, Lal PG, Warren BA, Xu Y, Honchell CD, Becha SD,
PI Lehr-Mason PM;
XX
DR WPI; 2003-636761/60.

DR N-PSDB; ADE79035.
XX
PT New human protein modification and maintenance molecules and
PT polynucleotides, useful for diagnosing, treating or preventing autoimmune
PT or inflammatory disorders (e.g. AIDS, allergy or anemia), multiple
PT sclerosis or cancer.
XX
PS Claim 1; SEQ ID NO 15; 405pp; English;
XX
CC This invention relates to novel isolated human proteins, which are human
CC protein modification and maintenance molecules (PMMW). The cellular
CC processes regulating modification and maintenance of protein molecules
CC coordinate their function, conformation, stabilisation and degradation.
CC Each of these processes is mediated by key enzymes or proteins such as
CC kinases, phosphatases, proteases, protease inhibitors, isomerases,
CC transferases and molecular chaperones. Compounds which modulate the
CC proteins of the invention may have anti-HIV, anti-allergic,
CC antiinflammatory, antianaemic, antiparkinsonian, nootropic,
CC anticonvulsant, antiarteriosclerotic, antiasthmatic, immunosuppressive,
CC antithyroid, cytostatic, hepatotropic, dermatological, antidiabetic,
CC nephrotropic, antigout, thymimetic, neuroprotective, osteopathic,
CC antiarthritic, antiparasitic, antihelminthic, antipsoriatic, uropathic,
CC ophthalmological, antirheumatic, haemostatic, antibacterial, virucide,
CC protozoacide or fungicide activities. The DNA sequence which encode the
CC proteins of the invention may be useful for gene therapy. The human
CC protein modification and maintenance molecules (PMMWs), the DNA sequences
CC which encode them and their modulating compounds are useful for
CC diagnosing, treating or preventing disorders associated with aberrant
CC expression of PMMW, particularly cell-proliferative disorders (for
CC example arteriosclerosis, hepatitis, polycythaemia vera, psoriasis,
CC primary thrombocytopaenia or cancer), developmental disorders (for
CC example anaemia or mental retardation), neurological disorders (for
CC example Alzheimer's disease, Parkinson's disease or epilepsy),
CC autoimmune/inflammatory disorders (for example AIDS, allergies, asthma,
CC autoimmune thyroiditis, Crohn's disease, diabetes mellitus,
CC glomerulonephritis, Goodpasture's syndrome, multiple sclerosis,
CC arthritis, osteoporosis, pancreatitis, Sjogren's syndrome) or microbial
CC infections. The present sequence is the amino acid sequence of a human
CC PMMW of the invention.
XX
SQ Sequence 493 AA;
Alignment Scores:
Pred. No.: 8.92e-153 Length: 493
Score: 1407.00 Matches: 270
Percent Similarity: 88.4% Conservative: 11
Best Local Similarity: 84.9% Mismatches: 19
Query Match: 65.7% Indels: 18
DB: 7 Gap#: 5
US-10-825-632-4 (1-1197) x ADE78977 (1-493)
Qy 3 TTTGAAGGACCAAGACTCCCTTTAGAGCATCACCTGCTAGCTAGTACGTACCTAAT 62
Db 186 PheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSerTyrValAsn 205
Qy 63 CCTGGAGAGGTCAGCAAGGCTGACTGACCTGGCTACTCACATTTCTGTGTCATGAGT 122
Db 206 ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCysIleSerGln 225
Qy 123 CACTGTGACTTCTTTTATAGTATAGTAACTACCAAGAAATCCACACTGTGTGTCCTT 182
Db 226 HisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeu 245
Qy 183 TACAGCTATCAGTCTCTGAAGTATGACCACTGCAACAAAGAAATTTTGGCCACC 242
Db 246 TyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPheThrPalatnr 265
Qy 243 ATTTTGGATTACAGAGGTCCTCTTCTCTGACTATCTCTCCAGAAATTTTCTCTTTGAA 302
Db 266 IleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIlePheSerPheGlu 285
Qy 303 AGTACTACTGGATTACATTGTTATGGATGCTCTACAGCCTCATGATCTACAGCCTGA 362

```
Db      286 SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly 305
Qy      363 AGAAATATCTACTGTGCTGTTCATATATATGGTGGTCTCAGGGTCAAAATAGAAATGAC 422
Db      306 LysLysTyrProThrValLeuPheIleTyrGlyProGlnValGlnLeu---ValAsn 324
Qy      423 GATCAGGTGAGAGGACTCCAATAT-----CTAGCTTCT---CGATATGAT 464
Db      325 AsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrVal 344
Qy      465 TTCATTGACTTATGAGTCGTGGGCATC-----CACGGCTGGTCTCATGGAGATACCTC 518
Db      345 ValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPhe 364
Qy      519 TCCTGATGGCATTAAATCAGAGGTCAGATATCTCAGGGTTGCTATTGCTGGGGCCCA 578
Db      365 LysTyrLys-----MetValAlaIleAlaGlyAlaPro 375
Qy      579 GTCACTCTGTGATCTTCTATGATACAGGATACACGGAACGTTATATGGTCACTCCCTGAC 638
Db      376 ValThrLeuTyrIlePheTyrAspThrGlyTyrThrGluArgTyrMetGlyHisProAsp 395
Qy      639 CAGATGACAGGCTATTACTTAGGATCTGTGGCCATGCACAGCAGAAAGTCCCTCT 698
Db      396 GlnAsnGlnGlnGlyTyrLeuGlySerValAlaMetGlnAlaGlnLysPheProSer 415
Qy      699 GAACCAATCGTTTACTGCTCTTACATGGTTCTCTGATGAGAAATGTCATTTGCAAT 758
Db      416 GluProAsnArgLeuLeuLeuLeuHisGlyPheLeuAspGluAsnValHisPheAlaHis 435
Qy      759 ACAGTATATTACTAGTCTTTTATGTAGGCTCGAAGCCATATGATTTACAGATCTAT 818
Db      436 ThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAspLeuGlnIleTyr 455
Qy      819 CTCTAGGAGAGACACAGCATAAGAGTTCTCTGAATCGGGAGAACATTATGAATGCACTTT 878
Db      456 ProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyrGluLeuHisLeu 475
Qy      879 TTGCATCTACTTCAAGAAACCTTGATCAGTATTGCTGCTCTAAAGTGATA 932
Db      476 LeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeuLysValIle 493

RESULT 15
ABU92026
ID      ABU92026 standard; protein; 831 AA.
AC
AC      ABU92026;
XX
XX      15-JUL-2003 (first entry)
DT
DT      Human protein modification and maintenance molecule-6 (PMMW-6).
XX
XX      Human; protein modification and maintenance molecule; PMMW; cancer;
KW      cell proliferation disorder; atherosclerosis; neurological disorder;
KW      epilepsy; Huntington's disease; stroke; immune disorder; allergy;
KW      inflammatory disorder; AIDS; developmental disorder; hypothyroidism;
KW      Cushing's syndrome; gastrointestinal disorder; epithelial disorder;
KW      infection; cytostatic; antiarteriosclerotic; anticonvulsant; nootropic;
KW      neuroprotective; cerebroprotective; anti-HIV; antiallergic; vulnerary;
KW      antiinflammatory; thyromimetic.
XX
XX      Homo sapiens.
OS
XX      WO2003031939-A2.
PN
XX      17-APR-2003.
PD
XX      11-OCT-2002; 2002WO-US032850.
PF
XX      12-OCT-2001; 2001US-0329689P.
PR      25-OCT-2001; 2001US-0335703P.
PR      09-NOV-2001; 2001US-0348887P.
```

```
PR      28-NOV-2001; 2001US-0334145P.
PR      06-DEC-2001; 2001US-0337451P.
PR      14-DEC-2001; 2001US-0340584P.
XX
XX      (INCY-) INCYTE GENOMICS INC.
XX
XX      Rankumar J, Gorvad AB, Baughn MR, Emerling BM, Yang J, Lee SY,
XX      Tran WK, Becha SD, Duggan BM, Lee EA, Griffin JA, Li JX,
XX      Sprague WW, Hafalia AJA, Chawla NK, Lehr-Mason PM, Kable AE, Yue H,
XX      Marquis JP, Yao MG, Richardson TW, Tang TY, Jin P, Chien D,
XX      Bhatia U, Burrill JD, Lee S, Blake JJ, Ho A, Zheng W;
XX
XX      WPI; 2003-430274/40.
XX      N-PSDB; ACA92421.
XX
XX      New human protein modification and maintenance molecules (PMMW), useful
XX      for diagnosing, treating and preventing diseases or conditions associated
XX      with the aberrant PMMW expression e.g. cancer, atherosclerosis, or
XX      infections.
XX
XX      Claim 1; Page 242-243; 31lpp; English.
XX
XX      The present invention relates to the isolation of human protein
XX      modification and maintenance molecules (PMMW) and the polynucleotide
XX      sequences encoding them. A total of 40 PMMW polypeptides (designated PMMW
XX      1 to PMMW-40) are disclosed. The sequences of the invention are useful
XX      for diagnosing a condition or disease associated with the expression of
XX      PMMW in a subject, preparing a polyclonal or monoclonal antibody, and
XX      generating an expression profile of a sample containing the
XX      polynucleotides. The diseases or conditions associated with decreased
XX      expression or overexpression of PMMW are cell proliferation disorders
XX      (e.g. cancer, atherosclerosis), neurological disorders (e.g. epilepsy,
XX      Huntington's disease, stroke), immune/inflammatory disorders, (e.g. AIDS,
XX      allergies), developmental disorders (e.g. hypothyroidism, Cushing's
XX      syndrome), gastrointestinal or epithelial disorders, and infections. The
XX      PMMW polypeptides or their fragments are useful in screening compounds
XX      for effectiveness as agonists or antagonists of the polypeptides, or in
XX      altering the expression of the target polynucleotide and compounds that
XX      specifically bind to, or modulate the activity of the polypeptide.
XX      CC ABU92021-ABU92060 represent the human PMMW polypeptides of the invention
XX      SQ Sequence 831 AA;
```

```
Alignment Scores:
Pred. No.: 1.16e-152 Length: 831
Score: 1407.00 Matches: 270
Percent Similarity: 88.4% Conservative: 11
Best Local Similarity: 84.9% Mismatches: 19
Query Match: 65.7% Indels: 18
DB: Gaps: 5
```

US-10-825-632-4 (1-1177) x ABU92026 (1-831)

```
Qy      3 TTTGAGGACCAAGAGACTCCCTTTAGAGCATCACCTGACGTAGTCAGTTAGTCAAT 62
Db      524 PheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSerTyrValAsn 543
Qy      63 CTCTGAGAGGTGACAGGCTGACTGACCGTGGTACTCACAATTTCTGTCGATCAGTCAG 122
Db      544 ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysIleSerGln 563
Qy      123 CACTGTGACTTCTTTATAGTATAGTATACACGAGAAATCCACACTGTGTCTCCCTT 182
Db      564 HisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeu 583
Qy      183 TACAAGCTATCAAGTCTCTGAAGATCACCCAACTTGCAGAAACAAAGGAATTTTGGCCACC 242
Db      584 TyrLysLeuSerSerProGluAspProThrCysLysThrLysGluPheTrpAlaThr 603
Qy      243 ATTTTGGATTACAGAGGTCCTCTCTCTGACTATATCTCTCCAGAAATTTTCTCTTTTGA 302
Db      604 IleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIlePheSerPheGlu 623
```

```
QY 303 A G T A C T A C T G G A T T T A C A T T G T A T G G A T C C T C A C A G C C T C A T G A T C T A C A G C C T G G A 362
Db |||||||
QY 624 S e r T h r G l y P h e T h r L e u T y r G l y M e t L e u T y r L y s P r o H i s A s p L e u G l n P r o G l y 643
Db |||||||
QY 363 A A G A A A T A T C C T A C T G T G C T T C A T A T A T G G T G G T C C T C A G G G T C A A A T A G A A A T T G A C 422
Db |||||||
QY 644 L y s L y s T y r P r o T h r V a l L e u P h e I l e T y r G l y P r o G l n V a l G l n L e u --- V a l A s n 662
Db |||||||
QY 423 G A T C A G G T G G A A G A C T C C A A T A T --- C T A G C T T C T --- C G A T A T G A T 464
Db |||||||
QY 663 A s n A r g P h e L y s G l y V a l L y s T y r P h e A r g L e u A s n T h r L e u A l a S e r L e u G l y T y r V a l 682
Db |||||||
QY 465 T T C A T T G A C T T A G A T C G T G T G G G A T C --- C A C G G T G G T C C T A T G G A G G A T A C C T C 518
Db |||||||
QY 683 V a l V a l V a l I l e A s p A s n A r g L y s e r C y s H i s A r g G l y L e u L y s P h e G l u G l y A l a P h e 702
Db |||||||
QY 519 T C C C T G A T G G C A T T A A T G C A G A G G T C A G A T A T C T T C A G G G T T G C T A T T G C T G G G C C C C A 578
Db |||||||
QY 703 L y s T y r L y s --- M e t V a l a l a l e A l a G l y A l a P r o 713
Db |||||||
QY 579 G T C A C T C T G T G G A T C T T C T A T G A T A C A G G A C G T T A T A T G G G T C A C C C T G A C 638
Db |||||||
QY 714 V a l T h r L e u T r p I l e P h e T y r A s p T h r G l y T y r T h r G l u A r g T y r M e t G l y H i s P r o A s p 733
Db |||||||
QY 639 C A G A A T G A A C A G G G C T A T T A C T T A G G A T C T G T G G C C A T C C A A G C A G A A A A G T T C C C T C T 698
Db |||||||
QY 734 G l n A s n G l u G l n G l y T y r T y r L e u G l y S e r V a l A l a M e t G l n A l a G l u L y s P h e P r o S e r 753
Db |||||||
QY 699 G A A C C A A T C G T T T A C T G C T T T A C A T G G T T C C T G G A T G A G A A T G C C A T T T T G C A C A T 758
Db |||||||
QY 754 G l u P r o A s n A r g L e u L e u L e u H i s G l y P h e L e u A s p G l u A s n V a l H i s P h e A l a H i s 773
Db |||||||
QY 759 A C C A G T A T A T T A C T G A G T T T T T A G T G A G G G C T G G A A G C C A T A T A T G A T T T A C A G A T C T A T 818
Db |||||||
QY 774 T h r S e r I l l e L e u L e u S e r P h e L e u V a l A r g A l a G l y L y s P r o T y r A s p L e u G l n L e u T y r 793
Db |||||||
QY 819 C C T C A G G A G A C A C A C A T A G A G T T C C T G A T C G G G A G A A C A T T A T G A A C T G A T C T T 878
Db |||||||
QY 794 P r o G l n G l u A r g H i s S e r I l l e A r g V a l P r o G l u S e r G l y G l u H i s T y r G l u L e u H i s L e u 813
Db |||||||
QY 879 T T G C A C T A C C T T C A A G A A A C C T T G G A T C A C G T A T T G C T G C T C T A A A A G T G A T A 932
Db |||||||
QY 814 L e u H i s T y r L e u G l n G l u A s n L e u G l y S e r A r g I l e A l a A l a L e u L y s V a l I l e 831
Db |||||||
```

Search completed: May 2, 2006, 01:03:09
Job time : 200.717 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: May 2, 2006, 02:50:19 ; Search time 4.00111 Seconds

(without alignments)

2721.092 Million cell updates/sec

Title: US-10-825-632-4

Perfect score: 2140

Sequence: 1 attttgaaggaccaaagac.....ttactaaaaaaaaaaaaa 1197

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 232119 seqs, 45477862 residues

Total number of hits satisfying chosen parameters: 464238

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DRV=xlp

-Q=/abs/ABSSWEB_spool/US10825632/runat_01052006_110000_3533/app_query.fasta.1

-DB=Published Applications_AA_New -QFMT=fastan -SUFFIX=n2p.rapbn -MINMATCH=0.1

-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62

-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100

-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0

-MAXLEN=200000000 -HOST=abs02p

-USER=US10825632 @CGN_1_1_59 @runat_01052006_110000_3533 -NCPUL=6 -ICPU=3

-NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120

-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7

-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications_AA_New.*

1: /SID55/ptodata/1/pubpaa/US08_NEW_PUB.pep1.*

2: /SID55/ptodata/1/pubpaa/US06_NEW_PUB.pep.*

3: /SID55/ptodata/1/pubpaa/US07_NEW_PUB.pep.*

4: /SID55/ptodata/1/pubpaa/US08_NEW_PUB.pep.*

5: /SID55/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*

6: /SID55/ptodata/1/pubpaa/US05_NEW_PUB.pep.*

7: /SID55/ptodata/1/pubpaa/US09_NEW_PUB.pep1.*

8: /SID55/ptodata/1/pubpaa/US10_NEW_PUB.pep.*

9: /SID55/ptodata/1/pubpaa/US10_NEW_PUB.pep1.*

10: /SID55/ptodata/1/pubpaa/US11_NEW_PUB.pep.*

11: /SID55/ptodata/1/pubpaa/US11_NEW_PUB.pep1.*

12: /SID55/ptodata/1/pubpaa/US60_NEW_PUB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	1645.5	76.9	882	11	US-11-151-601-20 Sequence 20, Appl
2	1093.5	51.1	497	11	US-11-264-096-1594 Sequence 1594, Ap
3	1089.5	50.9	497	11	US-11-264-096-1593 Sequence 1593, Ap
4	577	27.0	109	11	US-11-176-951-10 Sequence 10, Appl
5	427.5	20.0	745	11	US-11-079-463-6408 Sequence 6408, Ap

6	394	18.4	99	11	US-11-176-951-11	Sequence 11, Appl
7	341.5	16.0	760	11	US-11-208-288-6	Sequence 6, Appl
8	341	15.9	627	11	US-11-079-463-7758	Sequence 7758, Ap
9	337.5	15.8	738	11	US-11-208-288-4	Sequence 4, Appl
10	337.5	15.8	766	9	US-10-501-035-234	Sequence 234, App
11	337.5	15.8	766	11	US-11-208-288-2	Sequence 2, Appl
12	333.5	15.6	762	11	US-11-116-939-13	Sequence 13, Appl
13	333.5	15.6	766	9	US-10-522-789-2	Sequence 2, Appl
14	302	14.1	760	11	US-11-186-284-55	Sequence 55, Appl
15	180.5	8.4	624	11	US-11-079-463-7504	Sequence 7504, Ap
16	164	7.7	349	11	US-11-264-096-1591	Sequence 1591, Ap
17	152.5	7.1	657	11	US-11-179-977-1	Sequence 1, Appl
18	145	6.8	102	11	US-11-176-951-7	Sequence 7, Appl
19	142.5	6.7	737	11	US-11-079-463-9281	Sequence 9281, Ap
20	139	6.5	115	11	US-11-176-951-9	Sequence 9, Appl
21	139	6.5	115	11	US-11-176-951-12	Sequence 12, Appl
22	133.5	6.2	102	11	US-11-176-951-8	Sequence 8, Appl
23	106.5	5.0	315	11	US-11-018-868-143	Sequence 143, App
24	105.5	4.9	352	11	US-11-098-686-11006	Sequence 11006, A
25	98.5	4.6	316	11	US-11-079-463-9481	Sequence 9481, Ap
26	98.5	4.6	668	9	US-10-454-437-118	Sequence 118, App
27	95	4.4	314	11	US-11-179-977-17	Sequence 17, Appl
28	93.5	4.4	256	11	US-11-179-977-2	Sequence 2, Appl
29	92	4.3	397	11	US-11-084-508-5	Sequence 5, Appl
30	91	4.3	507	11	US-11-188-298-5830	Sequence 5830, Ap
31	88.5	4.1	255	11	US-11-179-977-3	Sequence 3, Appl
32	88	4.1	342	11	US-11-087-099-5518	Sequence 5518, Ap
33	88	4.1	1147	9	US-10-453-372-4	Sequence 4, Appl
34	88	4.1	1179	9	US-10-912-580-1	Sequence 1, Appl
35	88	4.1	1181	9	US-10-912-580-2	Sequence 2, Appl
36	88	4.1	1181	9	US-10-912-580-3	Sequence 3, Appl
37	87	4.1	308	11	US-11-098-568A-32737	Sequence 32737, A
38	84.5	3.9	309	11	US-11-098-568A-32738	Sequence 32738, A
39	84.5	3.9	335	11	US-11-098-568A-32737	Sequence 32737, A
40	83.5	3.9	344	11	US-11-079-463-7255	Sequence 7255, Ap
41	83.5	3.9	420	11	US-11-079-463-6314	Sequence 6314, Ap
42	83.5	3.9	2323	9	US-10-793-626-760	Sequence 760, App
43	83.5	3.9	3130	9	US-10-453-372-42	Sequence 42, Appl
44	83.5	3.9	3483	9	US-10-453-372-40	Sequence 40, Appl
45	83.5	3.9	3546	9	US-10-453-372-32	Sequence 32, Appl

ALIGNMENTS

RESULT 1
US-11-151-601-20
; Sequence 20, Application US/11151601
; Publication No. US20060003413A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Olandt, Peter J.
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Curtis, Rory A. J.
; APPLICANT: Williamson, Mark
; APPLICANT: Welch, Nadine
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASE, PHOSPHATASE.
; TITLE OF INVENTION: AND PROTEASE FAMILY MEMBERS AND USES THEREOF
; FILE REFERENCE: ME100-054PFCPIOMIDVIM
; CURRENT APPLICATION NUMBER: US/11/151,601
; CURRENT FILING DATE: 2005-06-13
; PRIOR APPLICATION NUMBER: US 10/170,789
; PRIOR FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: US 09/797,039
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: PCT/US01/06529
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/186,061
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 09/882,166
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: PCT/US01/19269
; PRIOR FILING DATE: 2001-06-15

```
; PRIOR APPLICATION NUMBER: US 60/212,078
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 09/934,406
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: PCT/US01/26052
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/226,740
; PRIOR FILING DATE: 2000-08-21
; Remaining prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 882
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-151-601-20

Alignment Scores:
Pred. No.: 1,648-160 Length: 882
Score: 1645.50 Matches: 310
Percent Similarity: 86.4% Conservative: 0
Best Local Similarity: 86.4% Mismatches: 0
Query Match: 76.9% Indels: 49
DB: 11 Gaps: 1

US-10-825-632-4 (1-1197) x US-11-151-601-20 (1-882)

Qy 3 TTGAAGGACCAAGATCCCTTTAGAGCATCACCTGTACGTAGTACGTAAAT 62
Db 524 PheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSerTyrValAsn 543
Qy 63 CTGGAGAGGTGACAGGCTGACTGACCGTGGCTACTCACATCTTGTGCTCATCAG 122
Db 544 ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysHisSerGln 563
Qy 123 CACTGTGACTCTTTTAAAGTAAATAGTAAACAGAGAAATCCACACTGTGTGCTCTT 182
Db 564 HisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeu 583
Qy 183 TACAAGCTATCAAGTCTCAAGATGACCACTTGGCAAAACAAAGGAATTTGGGCCACC 242
Db 584 TyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPheTAlaThr 603
Qy 243 ATTTTGGATTGACAGGTCCTCTCTGACTATATCTCTCCAGAAATTTCTCTTTGAA 302
Db 604 IleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePheSerPheGlu 623
Qy 303 AGTACTACTGGATTACATTGTGGATGCTCTACAGCTCATGATCTACAGCCTGGA 362
Db 624 SerThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly 643
Qy 363 AAGAAATATCTACTGCTGCTTCATATATATGCTGCTCTCAG----- 404
Db 644 LysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeuValAsnAsn 663
Qy 404 ----- 404
Db 664 ArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrValVal 683
Qy 404 ----- 404
Db 684 ValValIleAspAsnArgLysSerCysHisArgGlyLeuLysPheGluGlyAlaPheLys 703
Qy 405 -----GGTCAATAGAAATTTGACGATCAGGTGGAAGGACCTCCAATATCTAGCTTCT 455
Db 704 TyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyrLeuAlaSer 723
Qy 456 CGATATGATTTCATTGACTTAGATCGTGTGGGCATCCACGGCTGGCTCTATGGAGGATAC 515
Db 724 ArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrSerTyrGlyGlyTyr 743
Qy 516 CTCTCCCTGATGCATTATGACAGAGGTGAGATATCTTCAGGGTTGCTATTGCTGGGGCC 575
Db
```

```
Db 744 LeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAlaGlyAla 763
Qy 576 CCAGTCACCTCTGTGGATCTTCTATGATACAGGATACACGGAACTTATATGGTFCACCT 635
Db 764 ProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMetGlyHisPro 783
Qy 636 GACCAGAATGAACAGGGCTATTACTTAGGATCTGTGGCCATGTGCAAGCAGAAAAGTTCC 695
Db 784 AspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGluLysPhePro 803
Qy 696 TCTGAACCAAAATCGTTTACTGCTCTTACATGGTTTCTTGGATGAGAAATGTCCTTTTGA 755
Db 804 SerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnValHisPheAla 823
Qy 756 CATACCAAGTATATTACTGAGTTTTTTAGTGAGGCTGGAAGCCATATGATTTACAGATC 815
Db 824 HisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAspLeuGlnIle 843
Qy 816 TATCCTCAGAGAGACACAGCATAGAGTCTCTGAATCGGGAGAACATTTATGAACATCAT 875
Db 844 TyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyrGluLeuHis 863
Qy 876 CTTTTCGACTACTTCAGAAACCTTCGATCAGTATTTGCTGCTCTTAAAAGTGATA 932
Db 864 LeuLeuHisTyrLeuGlnGlnAsnLeuGlySerArgIleAlaLeuLysValIle 882

RESULT 2
US-11-264-096-1594
; Sequence 1594, Application US/11264096
; Publication No. US20060084794A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF546D1
; CURRENT APPLICATION NUMBER: US/11/264,096
; PRIOR FILING DATE: 2005-11-02
; PRIOR APPLICATION NUMBER: 09/833,245
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 2267
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1594
; LENGTH: 497
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-264-096-1594

Alignment Scores:
Pred. No.: 6,776-104 Length: 497
Score: 1093.50 Matches: 200
Percent Similarity: 69.3% Conservative: 42
Best Local Similarity: 57.3% Mismatches: 58
Query Match: 51.1% Indels: 49
DB: 11 Gaps: 1

US-10-825-632-4 (1-1197) x US-11-264-096-1594 (1-497)

Qy 3 TTGAAGGACCAAGATCCCTTTAGAGCATCACCTGTAGTACGTAGTACGTAAAT 62
Db 149 PheGlnGlyThrLysAspThrProLeuGluHisLeuTyrValValSerTyrGluAla 169
Qy 63 CTGGAGAGGTGACAAAGGCTGACTGACCGTGGCTACTCACATTTCTGCTGCATCAGTAC 122
Db 169 AlaGlyGluIleValArgLeuThrThrProGlyPheHisSerCysSerMetSerGln 189
Qy 123 CACTGTGACTCTTTTATAAGTAAATAGTAAACCAAGAAATCCACACTGTGTGCTCCTT 182
Db 189 AsnPheAspMetPheValSerHisTyrSerValSerThrProProCysValHisVal 208
```



```
QY 183 TACAAGCTATCAAGTCTCGAAGATGACCCAACTTGCAGAAACAAAGAAATTTGGCCACC 242
Db 209 TyrLysLeuSerGlyProAspAspProLeuHisLysGlnProArgPheTrpAlaSer 228
QY 243 ATTTGGATTACAGAGTCTCTCTCTGACTACTCTCCAGAAATTTCTCTTTTGA 302
Db 229 MetMetGluAlaAlaSerCysProProAspTyrValProGluLeuPheHisPheHis 248
QY 303 AGTACTACTGGATTACATTGTATGGGATGCTCTCAAGCCTCATGATCTACAGCCTGGA 362
Db 249 ThrArgSerAspValArgLeuTyrGlyMetIleTyrLysProHisAlaLeuGlnProGly 268
QY 363 AAGAAATATCTACTGCTGTCTCATATATGCTGCTCTCAG----- 404
Db 269 LysLysHisProThrValLeuPheValTyrGlyProGlnValGlnLeuValAsnAsn 288
QY 404 ----- 404
Db 289 SerPheLysGlyIleLysTyrLeuArgLeuLeuAsnThrLeuAlaSerLeuGlyTyrAlaVal 308
QY 404 ----- 404
Db 309 ValValIleAspGlyArgGlySerCysGlnArgGlyLeuArgPheGluGlyAlaLeuLys 328
QY 405 -----CGTCAATAGAAATTTGACGATCAGGTCAGGAGGACTCCAATATCTAGCTTCT 455
Db 329 AsnGlnMetGlyGlnValGluLeuGluAspGlnValGluGlyLeuGlnPheValAlaGlu 348
QY 456 CDATATGATTTCATTGACTTAGATCGTGTGGGCATCCAGCGTGTCTATGAGGATAC 515
Db 349 LysTyrGlyPheIleAspLeuSerArgValAlaIleHisGlyTrpSerTyrGlyPhe 368
QY 516 CTCTCCCTGATGCAATATGACAGGTCAGATATCTTCAGGGTGTCTATGCTGGGGCC 575
Db 369 LeuSerLeuMetGlyLeuIleHisLysProGlnValPheLysValAlaIleAlaGlyAla 388
QY 576 CAGTCACTCTGGGATCTTCATGATACAGGATACAGGACGTTATATGGTCAACCT 635
Db 389 ProValThrValTrpMetAlaTyrAspThrGlyTyrThrGluArgTyrMetAspValPro 408
QY 636 GACCAGAAATCAACAGCGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAAGTCTCCC 695
Db 409 GluAsnAsnGlnHisGlyTyrGluAlaGlySerValAlaLeuHisValGluLysLeuPro 428
QY 696 TCTGAACCAATCGTTTACTGCTTACATGCTTCTGATGCTGATGAGAAATGTCATTTGCA 755
Db 429 AsnGluProAsnArgLeuLeuIleLeuHisGlyPheLeuAspGluAsnValHisPhePhe 448
QY 756 CATACCAATATATCTAGTATTTTGTAGGGCTGGAAGCCATATGATTTACAGATC 815
Db 449 HisThrAsnPheLeuValSerGlnLeuIleArgAlaGlyLysProTyrGlnLeuGlnIle 468
QY 816 TATCTCAGAGAGACACAGCATAGAGTTCTCTGATCGGAGACATATGACTGCAT 875
Db 469 TyrProAsnGluArgHisSerIleArgCysProGluSerGlyGluHisTyrGluValThr 488
QY 876 CTTTGTGCACTACCTTCAAGAAACCTT 902
Db 489 LeuLeuHisPheLeuGlnGluTyrLeu 497
```

RESULT 3
US-11-264-096-1593
; Sequence 1593, Application US/11264096
; Publication No. US20060084794A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF546D1
; CURRENT APPLICATION NUMBER: US/11/264,096
; CURRENT FILING DATE: 2005-11-02
; PRIOR APPLICATION NUMBER: 09/833,245
; PRIOR FILING DATE: 2001-04-12

; PRIOR APPLICATION NUMBER: 60/229, 358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256, 931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199, 384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 2267
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1593
; LENGTH: 497
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (183)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-11-264-096-1593

Alignment Scores:
Pred. No.: 1,74e-103 Length: 497
Score: 1089.50 Matches: 199
Percent Similarity: 69.1% Conservative: 42
Best Local Similarity: 57.0% Mismatches: 59
Query Match: 50.9% Indels: 49
DB: 11 Gaps: 1

US-10-825-632-4 (1-1197) x US-11-264-096-1593 (1-497)

QY 3 TTTCAGGCGCCAAAGACTCCCTTTAGAGCATCACTGTAGCTAGTCAGTTACGTAAT 62
Db 149 PheGlnGlyThrLysAspThrProLeuGluHisLysLeuTyrValIleSerTyrGluAla 168
QY 63 CTGCGAGGTGACAGGCTGACTGACCGTGGCTACTCATTCTGTGTCATCAGTCAG 122
Db 169 AlaGlyGluIleValArgLeuThrThrProGlyPheSerHis***CysSerMetSerGln 188
QY 123 CACTGTGACTCTTTTATAGTAACTAGTAACTAGTAACTAGTAACTAGTAACTAGTAACT 182
Db 189 AsnPheAspMetPheValSerHisTyrSerValSerValSerValSerValHisVal 208
QY 183 TACAAGCTATCAAGTCTCGAAGATGACCCAACTTGCAGAAACAAAGAAATTTGGCCACC 242
Db 209 TyrLysLeuSerGlyProAspAspProLeuHisLysGlnProArgPheTrpAlaSer 228
QY 243 ATTTGGATTACAGAGTCTCTCTCTGACTACTCTCCAGAAATTTCTCTTTTGA 302
Db 229 MetMetGluAlaAlaSerCysProProAspTyrValProGluLeuPheHisPheHis 248
QY 303 AGTACTACTGGATTACATTGTATGGGATGCTCTCAAGCCTCATGATCTACAGCCTGGA 362
Db 249 ThrArgSerAspValArgLeuTyrGlyMetIleTyrLysProHisAlaLeuGlnProGly 268
QY 363 AAGAAATATCTACTGCTGTCTCATATATGCTGCTCTCAG----- 404
Db 269 LysLysHisProThrValLeuPheValTyrGlyProGlnValGlnLeuValAsnAsn 288
QY 404 ----- 404
Db 289 SerPheLysGlyIleLysTyrLeuArgLeuLeuAsnThrLeuAlaSerLeuGlyTyrAlaVal 308
QY 404 ----- 404
Db 309 ValValIleAspGlyArgGlySerCysGlnArgGlyLeuArgPheGluGlyAlaLeuLys 328
QY 405 -----CGTCAATAGAAATTTGACGATCAGGTCAGGAGGACTCCAATATCTAGCTTCT 455
Db 329 AsnGlnMetGlyGlnValGluLeuGluAspGlnValGluGlyLeuGlnPheValAlaGlu 348
QY 456 CGATATGATTTCATTGACTTAGATCGTGTGGGCATCCAGCGTGTCTATGAGGATAC 515
Db 349 LysTyrGlyPheIleAspLeuSerArgValAlaIleHisGlyTrpSerTyrGlyPhe 368
QY 516 CTCTCCCTGATGCAATATGACAGGTCAGATATCTTCAGGGTGTCTATGCTGGGGCC 575

```
||||| 369 LeuSerLeuMetGlyLeuIleHleHisProGlnValPheLysValAlaIleAlaGlyAla 388
||||| 576 CAGTCACCTCTGTCATCTCTATGATACAGGATACACGGAAGCTTATATGGCTCACCT 635
||||| 389 ProValThrValTrpMetAlaTyrAspThrGlyTyrThrGluArgTyrMetAspValPro 408
||||| 636 GACCAGAAATGAACAGGCTATTACTTAGGATCTCTGGCCATGCAAGCAGAAAGTTCCCC 695
||||| 409 GluAsnAsnGlnHisGlyTyrGluAlaGlySerValAlaLeuHisValGluLysLeuPro 428
||||| 696 TCTGAACCAAAATCGTTTACTCTCTTACATGTTTCTCTGATGAGAAATGCCATTTTGA 755
||||| 429 AsnGluProAsnArgLeuLeuIleLeuHisGlyPheLeuAspGluAsnValHisPhePhe 448
||||| 756 CATACAGTATATTACTAGTGTCTTCTAGTGGGCTGGAAGCCATATGATTTACAGATC 815
||||| 449 HisThrAsnPheLeuValSerGlnLeuIleArgAlaGlyLysProTyrGlnLeuGlnIle 468
||||| 816 TATCCTCAGGAGACAGCATAGAGTTCTCTGAATCGGAGAACATATTGACTGCAT 875
||||| 469 TyrProAsnGluArgHisSerIleArgCysProGluSerGlyGluHisTyrGluValThr 488
||||| 876 CTTTGCACCTACTCTCAAGAAACCTT 902
||||| 489 LeuLeuHisPheLeuGlnGluTyrLeu 497
```

RESULT 4

```
US-11-176-951-10
; Sequence 10, Application US/11176951
; Publication No. US20060024313A1
; GENERAL INFORMATION:
; APPLICANT: CHEN, YUAN
; APPLICANT: CHEN, YUAN-SHOU
; TITLE OF INVENTION: AGENTS THAT DISRUPT DIMER FORMATION IN DPP-IV FAMILY OF
; FILE REFERENCE: PROLYL DIPEPTIDASES
; CURRENT APPLICATION NUMBER: US/11/176,951
; PRIOR FILING DATE: 2005-07-06
; PRIOR APPLICATION NUMBER: 60/586,095
; PRIOR FILING DATE: 2004-07-06
; PRIOR APPLICATION NUMBER: 60/585,952
; PRIOR FILING DATE: 2004-07-06
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 10
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-176-951-10
```

```
Alignment Scores:
Pred. No.: 5,128-51 Length: 109
Score: 577.00 Matches: 109
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 27.0% Indels: 0
DB: 11 Gaps: 0
```

US-10-825-632-4 (1-1197) x US-11-176-951-10 (1-109)

```
Qy 606 GATACCGGAACGTTATATGGTCACTCCCTGACCAAGATGACAGGGCTATTACTAGGA 665
Db 1 GlyTyrThrGluArgTyrMetGlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGly 20
Qy 666 TCTGTGGCATCAACGAGAAAGTTCCCTCTCTGAAACCAATCGTTTACTGCTTACAT 725
Db 21 SerValAlaMetGlnAlaGluLysPheProSerGluProAsnArgLeuLeuLeuHis 40
Qy 726 GGTTCCTCGATGAGAATGTCCATTTTGCATACACCATGATATATTACTGAGTTTTAGTG 785
Db 41 GlyPheLeuAspGluAsnValHisPheAlaHisThrSerIleLeuLeuSerPheLeuVal 60
```

```
Qy 786 AGGCTGGAAGAACCCATATGATTTACAGATCTATCTCTCAGGAGACACAGCATAGAGTT 845
Db 61 ArgAlaGlyLysProTyrAspLeuGlnIleTyrProGlnGluArgHisSerIleArgVal 80
Qy 846 CTTGAATCGGAGAACATTATGAACCTGCATCTTTTGCACCTACCTTCAAGAAACCTTGA 905
Db 81 ProGluSerGlyGluHisTyrGluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGly 100
Qy 906 TCACGTATTGCTGCTCTAAAGTGATA 932
Db 101 SerArgIleAlaAlaLeuLysValIle 109
RESULT 5
US-11-079-463-6408
; Sequence 6408, Application US/11079463
; Publication No. US20060073161A1
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTEROIDES FRAC
; FILE REFERENCE: PATH00-03DIV2
; CURRENT APPLICATION NUMBER: US/11/079,463
; CURRENT FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/128,705
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: US 09/540,209
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 10444
; SEQ ID NO 6408
; LENGTH: 745
; TYPE: PRT
; ORGANISM: B.fragilis
US-11-079-463-6408
```

```
Alignment Scores:
Pred. No.: 1,76e-35 Length: 745
Score: 427.50 Matches: 97
Percent Similarity: 46.8% Conservative: 66
Best Local Similarity: 27.9% Mismatches: 124
Query Match: 20.0% Indels: 61
DB: 11 Gaps: 6
```

US-10-825-632-4 (1-1197) x US-11-079-463-6408 (1-745)

```
Qy 3 TTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACCTGTACGTAGTCAGTTACGTAAT 62
Db 411 TyrThrSerAsnGluGluSerProMetArgGlnAlaValTyrLysIleAsp-----Arg 428
Qy 63 CTGGAGAGGTGACAGGCTGACTGACCGTGGCTACTCATTCTTGTGTCAGTCAGT 122
Db 429 LysGlyLysLysMetLysLeuSerAsnGlnProGlyThrAsnSerProIlePheSer 448
Qy 123 CACTGTGACTCTTTTATAGTATAGTAACCAAGAAATCCACACTGTGTGTCCTT 182
Db 449 SerMetLysTyrPheMetAsnLysPheThrSerLeuAspThrProMetLeuIleThrLeu 468
Qy 183 TACAAGCTATCAAGTCTCTGAAGATGACCCCACTTGCRAAACAAAGGAATTTTGGCCACC 242
Db 469 -----AsnAspAsnThrGlyLysValLeuLysThrLeuValThr 481
Qy 243 ATTTTGATTCACGAGTCTCTCTGACTACTACTCTCCA-----GAAATTTTCTCT 296
Db 482 AsnAspLysLeuLysGlnLysLeuAlaGluTyrAlaIleProGlnLysGluPhePheThr 501
Qy 297 TTTGAAAGTACTACTGGATTTTACATTTATGGATGCTCTACAGCCCTCATGATCTACAG 356
Db 502 PheLysThrThrGluGlyValAspLeuAsnGlyTyrMetLysProValAsnPheAsp 521
Qy 357 CTGGAAAGAAATATCTACTGCTGTTCATATATGTTGCTCTCAGGGTCAATAGAA 416
Db 522 ProAlaLysArgTyrProValLeuMetPheGlnTyrSerGlyProGlySerGlnGlnVal 541
Qy 417 ATTGAC----- 422
```

```

: f: |||
Db 542 LeuAspLysTrpGlyIleSerTrpGluThrTyrMetAlaSerLeuGlyTyrValValAla 561
422 -----
Db 562 CysValAspGlyArgGlyThrGlyGlyArgGlySerGluPheGlnLysCysThrTyrLeu 581
423 -----GATCAGGTGGAGGACTCCCAATATCTAGCTTCGCGA 458
582 AsnLeuGlyValLysGluAlaLysAspGlnValGluAlaLaLysTyrLeuGlyGly--- 600
459 TATGATTTCATTCAGCTAGATCGTGTGGGCATCCACGGCTGCTCTATGGAGGATACCTC 518
601 LeuProTyrValAspLysGlyArgIleGlyIleTrpGlyTrpSerPheGlyGlyTyrMet 620
519 TCCTGATGCGCTTAATGATGAGGTCCAGATATCTTCAGGGTGTGCTATTGCTGGGGCCCA 578
621 ThrIleMetSerMetSerGluGlyThrProValPheLysAlaGlyValAlaValAlaAla 640
579 GTCACTCTGTGGATCTTCTATGATACAGGATACAGGAACGTTATATGGGTCCACCTGAC 638
641 ProThrAspTrpLysTyrTyrAspThrValTyrThrGluArgPheMetArgThrProLys 660
639 CAGAATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAGTTCCCTCT 698
661 GluAsnAlaGluGlyTyrLysAlaAlaSerAlaPheSerArgAlaAspAsnLeu----- 678
699 GAACCAATCGTTTACTGCTTTACATGTTCTCGATGAGATGCAATGTCCATTTTGACAT 758
679 ---HisGlyAsnLeuLeuLeuValHisGlyMetAlaAspAsnValHisPheGlnAsn 697
759 ACCAGTATATTACTAGTTTTTTAGTGAGGGCTGGAAGCCATATGATTTACAGATCTAT 818
698 CysThrGluTyrAlaGluHisLeuValGlnLeuGlyLysGlnPheAspMetGlnValTyr 717
819 CCTCAGGAGACACAGCATAGAGTTCTCGATCGGGAGAACCATATTGAAGTGCATCTT 878
718 ThrAsnArgAsnHisSerIleTyrGlyGlyAsnThrArgAsnHisLeuTyrThrLysLeu 737
879 TTGCACCTACTCTCAAGAAACCTT 902
738 ThrAsnPhePheArgAsnAsnLeu 745

RESULT 6
US-11-176-951-11
; Sequence 11, Application US/11176951
; Publication No. US20060024313A1
; GENERAL INFORMATION:
; APPLICANT: CHEN, XIN
; TITLE OF INVENTION: AGENTS THAT DISRUPT DIMER FORMATION IN DPP-IV FAMILY OF
; FILE REFERENCE: PROLYL DIPEPTIDASES
; CURRENT APPLICATION NUMBER: US/11/176,951
; PRIOR FILING DATE: 2005-07-06
; PRIOR APPLICATION NUMBER: 60/586,095
; PRIOR FILING DATE: 2004-07-06
; PRIOR APPLICATION NUMBER: 60/585,952
; PRIOR FILING DATE: 2004-07-06
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 11
; LENGTH: 99
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-176-951-11

Alignment Scores:
Pred. No.: 3,08e-32 Length: 99
Score: 394.00 Matches: 72
Percent Similarity: 82.8% Conservative: 10
Best Local Similarity: 72.7% Mismatches: 17
Query Match: 18.4% Indels: 0
```

```

DB: 11 Gap#: 0
US-10-825-632-4 (1-1197) x US-11-176-951-11 (1-99)
QY 606 GGATACACGGAACGTTATATGGTCCACCTGACCAAGTAAACAGGCGCTATTACTTAGGA 665
DB 1 GlyTyrThrGluArgTyrMetAspValProGluAsnAsnGlnHisGlyTyrGluAlaGly 20
QY 666 TCTGTGCCCATGCAAGCAAGAAAGTTCCCTCTCGAACCAAAATCGTTTACTGCTCTTACAT 725
DB 21 SerValAlaLeuHisValGluLysLeuProAsnGluProAsnArgLeuLeuLeuHis 40
QY 726 GGTTCCTCGATGATGAATGTCCATTTTGCATACACATGATATATTACTGAGTTTTTAGTG 785
DB 41 GlyPheLeuAspGluAsnValHisPhePheHisThrAsnPheLeuValSerGlnLeuIle 60
QY 786 AGGCTCGAAAGCCATATGATTTACAGATCTATCCTCAGGAGACACAGCAGATAAGAGTT 845
DB 61 ArgAlaGlyLysProTyrGlnLeuGlnIleTyrProAsnGluArgHisSerIleArgCys 80
QY 846 CCTGAATCGGAGAACATTTATGAATGTCATCTTTTGCATCTTCAAGAAACCTT 902
DB 81 ProGluSerGlyGluHisTyrGluValThrLeuLeuHisPheLeuGlnGluTyrLeu 99

RESULT 7
US-11-208-288-6
; Sequence 6, Application US/11208288
; Publication No. US20060051366A1
; GENERAL INFORMATION:
; APPLICANT: CHANG, Chiwen
; TITLE OF INVENTION: USE OF SOLUBLE CD26 AS INHIBITOR OF
; FILE REFERENCE: ANGIOGENESIS AND INFLAMMATION
; CURRENT APPLICATION NUMBER: US/11/208,288
; CURRENT FILING DATE: 2005-08-18
; PRIOR APPLICATION NUMBER: US 60/605,013
; PRIOR FILING DATE: 2004-08-26
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 760
; TYPE: PRT
; ORGANISM: Mus musculus
US-11-208-288-6

Alignment Scores:
Pred. No.: 1.19e-26 Length: 760
Score: 341.50 Matches: 101
Percent Similarity: 42.1% Conservative: 51
Best Local Similarity: 28.0% Mismatches: 120
Query Match: 16.0% Indels: 89
DB: 11 Gaps: 14

US-10-825-632-4 (1-1197) x US-11-208-288-6 (1-760)
QY 15 AAAGACTCCCTTTAGAGCATCACCTGTACGTAGTCAGTTACGTAAATCCTGGAGAGTG 74
DB 417 LysGluMetProGlyGlyArgAsnLeuTyrLysIle----- 428
QY 75 ACAAGGCTGACTACCGTGGCTACTCACATCTTCTGTCATCAGT----- 119
DB 429 ---GlnLeuThrAsp-----HisThrAsnValLysCysLeuSerCysAspLeuAsnPro 445
QY 120 CAGCATGTGACTTCTTTATAAGTAAGTATAGTACCAGAAAGATCCACAC----- 170
DB 446 GluArgCysGlnTyrTyrAlaValSerPheSerLysGluAlaLysTyrTyrGlnLeuGly 465
QY 171 TGT-----GTGTCCCTTTTACAAGCTATCAAGTCCCTGAAGATGACCCCACTTGC 218
DB 466 CysTrpGlyProGlyLeuProLeuTyrThrLeuHisArgSerThrAspHisLysGluLeu 485
QY 219 AAAACAAGGAATTTGGGGCCACCATTTTGGATTTCAGCAGGT-----CCTCTTCT 269
DB 219 AAAACAAGGAATTTGGGGCCACCATTTTGGATTTCAGCAGGT-----CCTCTTCT 269
```

Db 486 ArgValLeuGlu-----AspIleSerAlaLeuAspArgMetLeuGln 499
QY 270 GACTATACCTCCGAGAA-----ATTTCTCTTTTGAAGTACTACTGGATTACA 320
Db 500 AspValGlnMetProSerLysLeuAspPheIleValLeuAsnGluThrArgPhe--- 518
QY 321 TTGTATGGGATCCTCTACAAGCCTCATGATCTACAGCCTGGAAGAAATATCTACTGTG 380
Db 519 TrpTyrGlnMetIleLeuProHis---PheAspLysSerLysLysTyrProLeuLeu 537
QY 381 CTGTTCATATATGGTGGCTCT-----PheAspLysSerLysLysTyrProLeuLeu 401
Db 538 LeuAspValTyrAlaGlyProCysSerGlnLysAlaAspAlaSerPheArgLeuAsnTrp 557
QY 401 ----- 401
Db 558 AlaThrTyrLeuAlaSerThrGluAsnIleIleValAlaSerPheAspGlyArgGlySer 577
QY 402 -----CAGGGTCAAAATAGAAAT 419
Db 578 GlyTyrGlnGlyAspLysIleMetHisAlaIleAsnArgArgLeuGlyThrLeuGluVal 597
QY 420 GAGCATCAGGTGGAAGGACTCCAAATATCTAGCTTCTCGATATGATTTTCATTGACTTAGAT 479
Db 598 GluAspGlnIleGluAlaArgGlnPheVal---LysMetGlyPheValAspSerLys 616
QY 480 CTGTGGGATCCAGCGTGTCTCTATGGAGGATACCTCTCCCTGATGGCATTAAATGCAG 539
Db 617 ArgValAlaIleTrpGlyTrpSerTyrGlyGlyThrValThrSerMetValLeuGlySer 636
QY 540 AGTCAGATATCTTCAGGTGTCTATCTCTGGGCCCCAGTCACTCTGGGATCTCTAT 599
Db 637 GlySerGlyValPheLysCysGlyIleAlaValAlaProValSerArgTrpGluTyrTr 656
QY 600 GATACAGGATACACGGAAGTATATGGT-----CACCTGACCAAGATGAACAGGCG 653
Db 657 AspSerValTyrThrGluArgTyrMetGlyLeuProIleProGluAspAsnLeuAspHis 676
QY 654 TATTACTTAGTCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTGCAACCAATCGTTTA 713
Db 677 TyrArgAsnSerThrValMetSerArgAlaGluHisPhe-----LysGlnValGluTyr 694
QY 714 CTGCTCTTACATGGTTCTCTGGATGAGAATGTCATTTTGCACATACCACTATATTACTG 773
Db 695 LeuLeuIleHisGlyThrAlaAspAspAsnValHisPheGlnGlnSerAlaGlnIleSer 714
QY 774 AGTTTCTTAGTGAGGCTGGAAGCCATATGATTTTACAGATCTATCTCTCAGGAGAGACAC 833
Db 715 LysAlaLeuValAspAlaGlyValAspPheGlnAlaMetTrpTyrThrAspGluAspHis 734
QY 834 AGCATAAAGATTCCTCGAATCGGAGAACATTATGAATGCTGCTTTTGCACCTACCTCAA 893
Db 735 GlyIleAlaSerSerThrAlaHisGlnHisIleTyrSerHisMetSerHisPheLeuGln 754
QY 894 GAA 896
Db 755 Gln 755

RESULT 8

US-11-079-463-7758
; Sequence 7758 Application US/11079463
; Publication No. US20060073161A1
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTERIOIDES FR
; FILE REFERENCE: FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/11/079,463
; CURRENT FILING DATE: 2005-03-14
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: US 09/540,209
; PRIOR FILING DATE: 2000-04-04

; NUMBER OF SEQ ID NOS: 10444
; SEQ ID NO 7758
; LENGTH: 627
; TYPE: PRT
; ORGANISM: B. fragilis
US-11-079-463-7758
Alignment Scores:
Pred. No.: 1,28e-26 Length: 627
Score: 341.00 Matches: 92
Percent Similarity: 43.2% Conservative: 60
Best Local Similarity: 26.1% Mismatches: 126
Query Match: 15.9% Indels: 74
DB: 11 Gaps: 9
US-10-825-632-4 (1-1197) x US-11-079-463-7758 (1-627)
QY 12 ACCAAAGACTCCCTTTAGAGCATCACCTGTAGCTAGTACGTAAATCTCTGGAGAG 71
Db 295 ThrGluPheSerProLeuGlnAsnAsnLeuPheArgLeuAsp---ThrLysThrGlyThr 313
QY 72 GTGACAAGGCTGACTGACCGTGGCTACTCACATTTCTGCTCATCAGTCAGTCAGCTGTGAC 131
Db 314 ArgThrProLeuGlySerAlaGluGlyValHisSerGlyGlnLeuSerProSerGlyArg 333
QY 132 TTCTTTTAAAGTAAGTATAGTAACACAGAAGATCCACACTGTGTGCTCCCTTTACAAGCTA 191
Db 334 TyrLeuIleAspGlnTyrAsnSerProThrValProArgSerIleAsnIleIleAspVal 353
QY 192 TCAAGTCTCGAAGATGACCACTTGCACAAAGGAATTTTGGGCCACCATTTGGAT 251
Db 354 GlnSerGlyLysSer-----ValAsnLeuLeuThr 363
QY 252 TCAGCAGGCTCTCTCTGACTATATCTCTCTCA-----GAAATTTCTCTTTTGAAGT 305
Db 364 AlaAlaAspProPheThrGlyTyrLysMetProGlyIleGluThrGlyThrIleLysAla 383
QY 306 ACTACTGGATTTACA---TTGTATGGGATGCTCTAAGCCTCATGATCTACAGCTCGGA 362
Db 384 AlaAspGlyLysThrAspLeuTyrTyrArgLeuIleLysProAlaAspPheAspProAsn 403
QY 363 AAGAAATATCTACTGCTGTCTTATATATGTTGCTCTCTCAGGGTCAATA----- 413
Db 404 LysLysTyrProAlaIleValTyrValTyrGlyGlyProHisAlaGlnLeuValThrAsn 423
QY 413 ----- 413
Db 424 GlyTrpGlnAsnGlyAlaArgGlyTrpAspIleTyrMetAlaAsnLysGlyTyrIleMet 443
QY 413 ----- 413
Db 444 PheThrValAspGlyArgGlySerSerAspArgGlyLeuAspPheGluAsnValThrPhe 463
QY 414 -----GAAATTCAGCATCAGGTGGAAGGACTCCAAATATCTAGCTTCT 455
Db 464 ArgGlnLeuGlyIleGluGlyArgAspGlnValLysGlyThrGluPheLeuLysSer 483
QY 456 CGATATGATTTCATTGACTTAGATCTGTGGGCATCCACGGCTGCTCTCTATGGAGGATAC 515
Db 484 ---LeuProTyrValAspGlyAsnArgIleGlyValHisGlyTrpSerPheGlyGlyHis 502
QY 516 CTCTCCCTGATGCATTATGAGAGGTGAGATATCTTTCAGGTTGCTTATTCCTGGGGCC 575
Db 503 MetThrAlaLeuLeuLeuArgTyrProGluIlePheLysValGlyValAlaGlyGly 522
QY 576 CGAGTCACTCTGTGGATCTTCTATATACAGGATACGGAACGTATATTCGGTCAACCT 635
Db 523 ProValIleAspTrpGlyTyrGlyValMetTyrGlyGluArgTyrMetAspThrPro 542
QY 636 GACCAGAATGAACAGGCTTATTACTTAGATCTGTGGCATCGAAGCAGAAAAGTTTCCC 695
Db 543 GlnSerAsnProLysGlyTyrLysGluCysAsnLeuLysAsnLeuAlaGlyAsnLeu---- 561


```
Query Match: 15.8% Indels: 75
DB: 9 Gaps: 13
US-10-825-632-4 (1-1197) x US-10-501-035-234 (1-766)
QY 15 AAAGACTCCCTTTAGAGCATCAGCTGACGTACGTACGTAAATCCTGGAGAGGTG 74
DB 423 LysGlyMetProGlyArgAsnLeuThrLysLeuLeuLeuLeuLeuLeuVal 442
QY 75 ACAAGGCTGACTGACCGTGGCTACTCATTCTGCTGTCATCAGT---CAGCAGCTGTGAC 131
DB 443 ThrCysLeu-----SerCysGluLeuAsnProGluArgCysGln 455
QY 132 TTCTTTATAGTAAGTATAGTAACACGAAGAATCCACAC-----TGT----- 173
DB 456 TyrTyrSerValSerPheSerLysGluAlaLysTyrTyrGlnLeuArgCysSerGlyPro 475
QY 174 ---GTGTCCCTTTACAGCTATCAAGTCTCCTGAAGATGACCCCACTTGCACAAAGGAA 230
DB 476 GlyLeuProLeuTyrThrLeuHisSerSerValAsnAspLysGlyLeuArgValLeuGlu 495
QY 231 TTTTGGGCCACCATTTGGATTGACGAGGTCTCTCTGACTATATCTCTCCAGAA--- 287
DB 496 Asp---AsnSerAlaLeuAspLysMet-----LeuGlnAsnValGlnMetProSerLys 512
QY 288 -----ATTTTCTCTTTGAAAGTACTACTGGATTACATTTGATGGGATGCTCTACAAG 341
DB 513 LysLeuAspPheIleLeuAsnGluThrLysPhe---TrpTyrGlnMetIleLeuPro 531
QY 342 CCTCATGATCTACAGCTCGAAAGAAATATCTACTGTGCTGTTCATATATGTTGCTCT 401
DB 532 ProHis---PheAspLysSerLysLysTyrProLeuLeuLeuAspValTyrAlaGlyPro 550
QY 401 ----- 401
DB 551 CysSerGlnLysAlaAspThrValPheArgLeuAsnTrpAlaThrTyrLeuAlaSerThr 570
QY 401 ----- 401
DB 571 GluAsnIleIleValAlaSerPheAspGlyArgGlySerGlyTyrGlnGlyAspLysIle 590
QY 402 -----CAGGTCACAAATAGAAATTTGACGATCAGGTGGAAGGACTC 440
DB 591 MethIleAlaIleAsnArgArgLeuGlyThrPheGluValGluAspGlnIleGluAlaAla 610
QY 441 CAATATCTAGCTTCTCGATATGATTTTCTGATCTTAGATCGTGTGGCCTCCAGGCTGG 500
DB 611 ArgGlnPhe---SerLysMetGlyPheValAspAsnLysArgIleAlaIleTrpGlyTrp 629
QY 501 TCCTATGGAGGATACCTCTCCCTGATGCGATTATGACAGGTGATGATATCTTCCAGGTT 560
DB 630 SerTyrGlyTyrValThrSerMetValLeuGlySerGlySerGlyValPheLysCys 649
QY 561 GCTATTGCTGGGCCCCAGTCTCTGTGGATCTTCTATGATACAGGATACAGGAACGT 620
DB 650 GlyIleAlaValAlaProValSerArgTrpGluTyrTyrAspSerValTyrThrGluArg 669
QY 621 TATATGGT-----CACCTGACCATGACAGGCTATTACTAGGATCTGTGGCC 674
DB 670 TyrMetGlyLeuProThrProGluAspAsnLeuAspHisTyrArgAsnSerThrValMet 689
QY 675 ATGACACGAGAAGTTCCTCTGACCAACCAATCGTTTACCTCTTACATGTTCTTCCCTG 734
DB 690 SerArgAlaGluAsnPhe-----LysGlnValGluTyrLeuLeuIleHisGlyThrAla 707
QY 735 GATGAGAAATGTCATTTTGCACATACCACTATATTTACTGAGTGTTTTATGAGGCTGGA 794
DB 708 AspAspAsnValHisPheGlnGlnSerAlaGlnIleSerLysAlaLeuValAspValGly 727
QY 795 AAGCCATATGATTACAGATCTATCTCTCAGGAGAGACACAGCATTAAGAGTTCCTGAATCG 854
DB 728 ValAspPheGlnAlaMetTrpTyrThrAspGluAspHisGlyIleAlaSerSerThrAla 747
```

```
QY 855 GGAGAACATTATGAACTGCATCTTTTGCACCTACCTTCAAGAA 896
DB 748 HisGlnHisIleTyrThrHisMetSerHisPheIleLysGln 761
RESULT 11
US-11-208-288-2
; Sequence 2, Application US/11208288
; Publication No. US20060051366A1
; GENERAL INFORMATION:
; APPLICANT: CHANG, Chiwen
; TITLE OF INVENTION: USE OF SOLUBLE CD26 AS INHIBITOR OF
; FILE REFERENCE: 39533-0001
; CURRENT APPLICATION NUMBER: US/11/208,288
; PRIOR FILING DATE: 2005-08-18
; PRIOR FILING DATE: 2004-08-26
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 766
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-208-288-2
Alignment Scores: 3.07e-26 Length: 766
Pred. No.: 337.50 Matches: 99
Score: Similarity: 42.9% Conservative: 53
Best Local Similarity: 28.0% Mismatches: 127
Query Match: 15.8% Indels: 75
DB: 11 Gaps: 13
US-10-825-632-4 (1-1197) x US-11-208-288-2 (1-766)
QY 15 AAAGACTCCCTTTAGAGCATCAGCTGACGTACGTACGTAAATCCTGGAGAGGTG 74
DB 423 LysGlyMetProGlyArgAsnLeuThrLysLeuLeuLeuLeuLeuVal 442
QY 75 ACAAGGCTGACTGACCGTGGCTACTCATTCTGCTGTCATCAGT---CAGCAGCTGTGAC 131
DB 443 ThrCysLeu-----SerCysGluLeuAsnProGluArgCysGln 455
QY 132 TTCTTTATAGTAAGTATAGTAACACGAAGAATCCACAC-----TGT----- 173
DB 456 TyrTyrSerValSerPheSerLysGluAlaLysTyrTyrGlnLeuArgCysSerGlyPro 475
QY 174 ---GTGTCCCTTTACAGCTATCAAGTCTCCTGAAGATGACCCCACTTGCACAAAGGAA 230
DB 476 GlyLeuProLeuTyrThrLeuHisSerSerValAsnAspLysGlyLeuArgValLeuGlu 495
QY 231 TTTTGGGCCACCATTTTGGATTGACGAGGTCTCTCTGACTATATCTCTCCAGAA--- 287
DB 496 Asp---AsnSerAlaLeuAspLysMet-----LeuGlnAsnValGlnMetProSerLys 512
QY 288 -----ATTTTCTCTTTGAAAGTACTACTGGATTACATTTGATGGGATGCTCTACAAG 341
DB 513 LysLeuAspPheIleLeuAsnGluThrLysPhe---TrpTyrGlnMetIleLeuPro 531
QY 342 CCTCATGATCTACAGCTCGAAAGAAATATCTACTGTGCTGTTCATATATGTTGCTCT 401
DB 532 ProHis---PheAspLysSerLysLysTyrProLeuLeuLeuAspValTyrAlaGlyPro 550
QY 401 ----- 401
DB 551 CysSerGlnLysAlaAspThrValPheArgLeuAsnTrpAlaThrTyrLeuAlaSerThr 570
QY 401 ----- 401
DB 571 GluAsnIleIleValAlaSerPheAspGlyArgGlySerGlyTyrGlnGlyAspLysIle 590
QY 402 -----CAGGTCACAAATAGAAATTTGACGATCAGGTGGAAGGACTC 440
```

Db 591 MetHisAlaIleAsnArgLeuGlyThrPheGluValGluAspGlnIleGluAlaA 610
Qy 441 CAATATCTAGCTTCTCGATATGATTTCAATTGATTCGTGGCGCATCCACGGCTGG 500
Db 611 ArgGlnPhe---SerLysMetGlyPheValAspAsnLysArgIleAlaIleTrpGlyTrp 629
Qy 501 TCGTATGGAGGATACCTCCCTGATGGCATTAATGCAGAGTGCAGATATCTCAGGGTT 560
Db 630 SerTyGlyGlyTyValThrSerMetValLeuGlySerGlySerGlyValPheLysCys 649
Qy 561 GCTATTGCTGGGCCCCAGCTACTCTGTGGATCTTCTATGATACAGGATCACCGAACGT 620
Db 650 GlyIleAlaValAlaProValSerArgTrpGluTyTyAspSerValTyThrGluArg 669
Qy 621 TATATGGGT-----CACCTCACCAGCAATGACAGGCTATTACTTAGGATCTGTGGCC 674
Db 670 TyrMetGlyLeuProThrProGluAspAsnLeuAspHisTyArgAsnSerThrValMet 689
Qy 675 ATCGAAGCGAAGATTCCTCTGACCAACCAATCGTTTACTGCTTACATGGTTTCTGTG 734
Db 690 SerArgAlaGluAsnPhe-----LysGlnValGluTyLeuLeuIleHisGlyThrAla 707
Qy 735 GATGAGAATGCTCATTTTGCATACACAGTATATTACTGAGTCTTTTGTAGTGAGGCTGGA 794
Db 708 AspAspAsnValHisPheGlnSerAlaGlnIleSerLysAlaLeuValAspValGly 727
Qy 795 AAGCCATATGATTTACAGATCTATCTCTCAGGAGACACACATAGAGTTCTCTGAATCG 854
Db 728 ValAspPheGlnAlaMetTrpTyThrAspGluAspHisGlyIleAlaSerSerThrAla 747
Qy 855 GAGAACATATGAACTGCATCTTTTGCACTACCTTCAAGAA 896
Db 748 HisGlnHisIleTyThrHisMetSerHisPheIleLysGln 761

RESULT 12

; Sequence 13, Application US/11116939
; Publication No. US20050265995A1
; GENERAL INFORMATION:
; APPLICANT: Stephen Tomlinson
; TITLE OF INVENTION: TISSUE TARGETED COMPLEMENT MODULATORS
; FILE REFERENCE: 19113.011502
; CURRENT APPLICATION NUMBER: US/11/116,939
; PRIOR FILING DATE: 2005-04-28
; PRIOR APPLICATION NUMBER: 60/565,907
; PRIOR FILING DATE: 2004-04-28
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 762
; TYPE: PRF
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence; note=synthetic
; OTHER INFORMATION: construct
US-11-116-939-13

Alignment Scores:
Pred. No.: 7,89e-26 Length: 762
Score: 333.50 Matches: 99
Percent Similarity: 42.2% Conservative: 52
Best Local Similarity: 27.7% Mismatches: 124
Query Match: 15.6% Indels: 83
Gaps: 14

US-10-825-632-4 (1-1197) x US-11-116-939-13 (1-762)

Qy 15 AAAGACTCCCTTTAGAGCATCACCTGTACGTACGTACGTAAATCCTGGAGAGGTG 74
Db 419 LysGlyMetProGlyGlyArgAsnLeuTyLysIle----- 430
Qy 75 ACAAGCTGACTGACCGGTGCTACTCACAATCTTGTGCTGATCAGT----- 119

RESULT 13

US-10-522-789-2
; Sequence 2, Application US/10522789
; Publication No. US20050260732A1
; GENERAL INFORMATION:
; APPLICANT: TANABE SEIYAKU CO., LTD.

Db 431 ---GlnLeuSerAsp-----TyrThrLysValThrCysLeuSerCysGluLeuAsnPro 447
Qy 120 CAGCACTGTCACATCTTTTATAAGTATAGTAACACAGAAATCCACAC----- 170
Db 448 GluArgCysGlnTyTySerValSerPheSerLysGluAlaLysTyTyGlnLeuArg 467
Qy 171 TGT-----GTGTCCTTTTCAAGCTATCAAGTCTCTGAAGATGACCAACTTGC 218
Db 468 CysSerGlyProGlyLeuProLeuTyThrLeuHisSerSerValAsnAspLysGlyLeu 487
Qy 219 AAAACAAGAAATTTTGGGCCACCATTTTGGATTTCAGCAGGCTCTCTCTCCTGACTATCT 278
Db 488 ArgValLeuGluAsp---AsnSerAlaLeuAspLysMet-----LeuGlnAsnValGln 504
Qy 279 CCTCCAGAA-----ATTTCCTCTTTTGAAGATCTACTTGGATTTCATTTGATGGG 329
Db 505 MetProSerLysLysLeuAspPheIleIleLeuAsnGluThrLysPhe---TrpTyGln 523
Qy 330 ATGCTCTACAAGCTCATGATCTACAGCTCGGAAGAAATATCTCTGCTGTGTTTCATA 389
Db 524 MetIleLeuProHis---PheAspLysSerLysTyTrpProLeuLeuAspVal 542
Qy 390 TATGCTGCTCCT----- 401
Db 543 TyrAlaGlyProCysSerGlnLysAlaAspThrValPheArgLeuAsnTrpAlaThrTyr 562
Qy 401 ----- 401
Db 563 LeuAlaSerThrGluAsnIleIleValAlaSerPheAspGlyArgGlySerGlyTyGln 582
Qy 402 -----CAGGTCAAATAGAAATTCACATCAG 428
Db 583 GlyAspLysIleMetHisAlaIleAsnArgArgLeuGlyThrPheGluValGluAspGln 602
Qy 429 GTGGAAGGACTCCCAATATCTAGCTTCTCGATATGATTTTCATTTGACTAGATCGTGGGC 488
Db 603 IleGluAlaAlaArgGlnPhe---SerLysMetGlyPheValAspAsnLysArgIleAla 621
Qy 489 ATCCACGCTGGTCTTATGGAGGATACCTCTCCCTGATGGCATTAATGCAGAGTTCAGAT 548
Db 622 IleTrpGlyTrpSerTyGlyTyValThrSerMetValLeuGlySerGlySerGly 641
Qy 549 ATCTTCAGGTTTCTATGCTGGGCCCCAGTCACTCTGTGATCTTCTATGATACAGA 608
Db 642 ValPheLysCysGlyIleAlaValAlaProValSerArgTrpGluTyTyAspSerVal 661
Qy 609 TACACGGAACGTTTATATGGT-----CACCTGACCAAGATGACACAGGCTATTACTTA 662
Db 662 TyrThrGluArgTyMetGlyLeuProThrProGluAspAsnLeuAspHisTyArgAsn 681
Qy 663 GGATCTGTGGCCATGCAAGCAGAAAGTTCCTCTCTGAACCAAAATCGTTTACTGCTCTTA 722
Db 682 SerThrValMetSerArgAlaGluAsnPhe-----LysGlnValGluTyLeuLeuIle 699
Qy 723 CATGGTTTCTGATGAGAATGTCATTTTGACATACACGATATATTAATGATTTTATA 782
Db 700 HisGlyThrAlaAspAspAsnValHisPheGlnGlnSerAlaGlnIleSerLysAlaLeu 719
Qy 783 GTGAGGCTCGAAAGCCATATGATTTACAGATCTATCTCAGGAGACACAGCATAGA 842
Db 720 ValAspValGlyValAspPheGlnAlaMetTrpTyThrAspGluAspHisGlyIleAla 739
Qy 843 GTTCCTGAATCGGAGAAACATTATGAACTGCTATTTTGCATCTACTCTCAAGAA 896
Db 740 SerSerThrAlaHisGlnHisIleTyThrHisMetSerHisPheIleLysGln 757

```
; TITLE OF INVENTION: Three-dimensional structure of dipeptidyl peptidase IV
; FILE REFERENCE: 03-039-PCT
; CURRENT APPLICATION NUMBER: US/10/522,789
; CURRENT FILING DATE: 2005-01-28
; PRIOR APPLICATION NUMBER: US 60/398,761
; PRIOR FILING DATE: 2002-07-29
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patent in version 3.1
; SEQ ID NO: 2
; LENGTH: 766
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-522-789-2

Alignment Scores:
Pred. No.: 7,9e-26 Length: 766
Score: 333.50 Matches: 99
Percent Similarity: 42.2% Conservative: 52
Best Local Similarity: 27.7% Mismatches: 124
Query Match: 15.6% Indels: 83
DB: 9 Gaps: 14

US-10-825-632-4 (1-1197) x US-10-522-789-2 (1-766)
QY 15 AAAGACTCCCTTTAGAGCATCACCTGTACGTACGTACGTAAATCTCGAGAGTG 74
Db 423 LysGlyMetProGlyGlyArgAsnLeuTyrLysIle-----434
QY 75 ACAAGGCTGACTGACCGTGGCTACTCACATCTTCTGCTGCATCAGT-----119
Db 435 ---GlnLeuSerAsp-----TyrThrLysValThrCysLeuSerCysGluLeuAenPro 451
QY 120 CAGCACTGTCATCTTTTATAAGTAAGTATAGTAAACAGAGATCCACAC-----170
Db 452 GluArgCysGlnTyrTyrSerValSerPheSerLysGluAlaLysTyrGlnLeuArg 471
QY 171 TGT-----GTGTCCTTTACAAGCTATCAAGTCTCTGAGATGCCCACTTGC 218
Db 472 CysSerGlyProGlyLeuProLeuTyrThrLeuHisSerValAsnAspLysGlyLeu 491
QY 219 AAACAAGAAATTTGGGCCACCAATTTGGATTTCAGCAGGTCTCTCTCGACTATAT 278
Db 492 ArgValLeuGluAsp---AsnSerAlaLeuAspLysMet-----LeuGlnAenValGln 508
QY 279 CCTCAGAA-----ATTTCCTCTTTGAAAGTACTACTGGAATTTACATTTGATGG 329
Db 509 MetProSerLysLysLeuAspPheIleLeuAenGluThrLysPhe---TrpTyrGln 527
QY 330 ATGCTCTACAAGCCTCATGATCTACAGCCTGGAAGAAATATCTACTGTGCTGTTCATA 389
Db 528 MetIleLeuProProHis---PheAspLysSerLysLysTyrProLeuLeuLeuAspVal 546
QY 390 TATGGTGTCTCT-----401
Db 547 TyrAlaGlyProCysSerGlnLysAlaAspThrValPheArgLeuAenTrpAlaThrTyr 566
QY 401 -----401
Db 567 LeuAlaSerThrGluAsnIleIleValAlaSerPheAspGlyArgGlySerGlyTyrGln 586
QY 402 -----CAGGTCAAATAGAAATGACGATCG 428
Db 587 GlyAspLysIleMetHisAlaIleAsnArgArgLeuGlyThrPheGluValGluAspGln 606
QY 429 GTGAAGAGACTCCAATATCTAGCTCTCGATATGATTTCAATTCATTCGTAGATCGTGGGC 488
Db 607 IleGluAlaArgGlnPhe---SerLysMetGlyPheValAspAsnLysArgIleAla 625
QY 489 ATCCAGGCTGCTCTATGGAGATACCTCTCCCTGATGGCATTAATGACAGGTGAGAT 548
Db 626 IleTrpGlyTrpSerTyrGlyGlyTyrValThrSerMetValLeuGlySerGlySerGly 645
QY 549 ATCTTCAGGGTTGCTATTGCTGGGGCCCCAGCTACTCTGTGGATCTTCTATGATACAGGA 608
```

```
Db 646 ValPheLysCysGlyIleAlaValAlaProValSerArgTyrGluTyrTyrAspSerVal 665
QY 609 TACACGGAACGTTATATGGT-----CACCTGACCAAGATGAACAGGGCTATTCTTA 662
Db 666 TyrThrGluArgTyrMetGlyLeuProThrProGluAspAenLeuAspHisTyrArgAsn 685
QY 663 GGATCTGTGCCATGCAAGCAGAAAGTTCCCTCTGAACCAATCGTTTACTGCTCTTA 722
Db 686 SerThrValMetSerArgAlaGluAenPhe-----LysGlnValGluTyrLeuLeuIle 703
QY 723 CATGGTTTCTCGATGAGATGTCATTTTGACACATACCATATATATATCTAGGTTTTTTA 782
Db 704 HisGlyThrAlaAspAspAenValHisPheGlnGlnSerAlaGlnIleSerLysAlaLeu 723
QY 783 GTCAGGCTGGAAGCCATATGATTTACAGATCTATCTCTCAGGAGACACAGCATAGA 842
Db 724 ValAspValGlyValAspPheGlnAlaMetTrpTyrThrAspGluAspHisGlyIleAla 743
QY 843 GTTCCTGAATCGGGAGAACATATGAACCTGTCATCTTTTGACACTACCTTCAAGAA 896
Db 744 SerSerThrAlaHisGlnHisIleTyrThrHisMetSerHisPheIleLysGln 761

RESULT 14
US-11-186-284-55
; Sequence 55, Application US/11186284
; Publication No. US20050266493A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Berger, Allison
; APPLICANT: Guillemette, Tracy L.
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Schlesinger, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Thibodeau, Stephen N.
; APPLICANT: Burgart, Lawrence J.
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; THERAPY OF COLON CANCER
; FILE REFERENCE: MEM01-029PZRNM
; CURRENT APPLICATION NUMBER: US/11/186,284
; CURRENT FILING DATE: 2005-07-21
; PRIOR APPLICATION NUMBER: US/10/301,822
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 60/339,971
; PRIOR FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: US 60/361,978
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/381,988
; PRIOR FILING DATE: 2002-05-20
; NUMBER OF SEQ ID NOS: 228
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 55
; LENGTH: 760
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-11-186-284-55

Alignment Scores:
Pred. No.: 1,35e-22 Length: 760
Score: 302.00 Matches: 93
Percent Similarity: 40.5% Conservative: 55
Best Local Similarity: 25.5% Mismatches: 139
Query Match: 14.1% Indels: 78
DB: 11 Gaps: 11

US-10-825-632-4 (1-1197) x US-11-186-284-55 (1-760)
QY 3 TTGAAGGACCAAGACTCCCTTTAGAGCATCACCTGTAGTACGTAGTACGTAGTAAT 62
Db 401 PheArgValThrGlnAspSerLeuPheTyrSerSerAenGluPheGluGluTyrProGly 420
QY 63 CCTGGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACATTCT---TGCTGCATCAGT 119
```



```
Db 421 ArgArgAsnIleTyrArgIleSerIleGlySerTyrProSerLysLysCysValThr 440
QY 120 CAGCAC-----TGTCACCTCTTTATAGTAAGTATAGTAACTAACACAGAGAT 164
Db 441 CysHisLeuArgLysGluArgCysGlnTyrThrAlaSerPheSerAspTyrAlaLys 460
QY 165 CCACACTGTGTCTCCCTTTAC-----AAGCTATCAAGTCTCTGAAGATGACCCA 212
Db 461 TyrTyrAlaLeuValCysTyrGlyProGlyIleProIleSerThrLeuHisAspGlyArg 480
QY 213 ACTGCAAAACAAAGAAATTTGGGCCACCAATTTGGATTGACAGAGTCTCTCTCTGAC 272
Db 481 ThrAspGlnGlu-----IleLysIleLeuGluGluAsnLysGluLeuGluAsn 496
QY 273 -----TATACCTCCAGAAATTTCTCTTTGAAAGTACTACTGGATT 317
Db 497 AlaLeuLysAsnIleGlnLeuProLysGluGluIleLysLysLeuGluValAspGluIle 516
QY 318 ACATTGTATGGGATCTCTACAAGCCTCATGATCTACAGCCTGGAAGAAATATCTACT 377
Db 517 ThrLeuTyrTyrLysMetIleLeuProGlnPheAspArgSerLysLysTyrProLeu 536
QY 378 GTGCTGTTCAATATGTTGGTCT-----GTCCT----- 401
Db 537 LeuIleGlnValTyrGlyGlyProCysSerGlnSerValArgSerValPheAlaValAsn 556
QY 401 ----- 401
Db 557 TrpIleSerTyrLeuAlaSerLysGluGlyMetValIleAlaLeuValAspGlyArgGly 576
QY 402 -----CAGGTCAAATAGAA 416
Db 577 ThrAlaPheGlnGlyAspLysLeuLeuTyrAlaValTyrArgLysLeuGluValTyrGlu 596
QY 417 ATGACGATCAGGTGAAGGACTCCATATCTAGCTCTCGATATGATTTTCATGACTTA 476
Db 597 ValGluAspGlnIleThrAlaValArgLysPheIle---GluMetGlyPheIleAspGlu 615
QY 477 GATCGTGTGGGATCCAGGCTGTCCTATCGAGATACCTCTCCCTGATGGCAATATG 536
Db 616 LysArgIleAlaIleTrpTyrSerTyrGlyTyrValSerSerLeuAlaLeuAla 635
QY 537 CAGAGTGCAGATATCTCAGGTTCTATTGCTGGGCCCCAGTCACCTCTGTGGATCTTC 596
Db 636 SerGlyThrGlyLeuPheLysCysGlyIleAlaValAlaProValSerSerTrpGluTyr 655
QY 597 TATGATACAGGATACACGGAAGTTATATGGTCCACCT-----GACCAGATGAACAG 650
Db 656 TyrAlaSerValTyrThrGluArgPheMetGlyLeuProThrLysAspAspAsnLeuGlu 675
QY 651 GdCTATTACTAGGATCTGTGGCCATGCAACGACGAAAGTTCCCTCTGAAACCAATCGT 710
Db 676 HisTyrLysAsnSerThrValMetAlaArgAlaGluTyrPheArgAsnValAsp----- 693
QY 711 TTACTGCTCTTACATGTTCTCTGATGAGAATGTTCATTTTGACATACACGATATATTA 770
Db 694 TyrLeuLeuIleHisGlyThrAlaAspAsnValHisPheGlnAsnSerAlaGlnIle 713
QY 771 CTGAGTTTTTATGAGGCTGGAAAGCCATATGATTTTACAGATCTATCTCTCAGGAGAGA 830
Db 714 AlaLysAlaLeuValAsnAlaGlnValAspPheGlnAlaMetTrpTyrSerAspGlnAsn 733
QY 831 CACAGCATAGAGTCTCTGAATCGGGA-----GAACATATGAACTGCATCTTTG 881
Db 734 HisGlyLeu-----SerGlyLeuSerThrAsnHisLeuTyrThrHisMetThr 749
QY 882 CACTACTCTTCAAGAA 896
Db 750 HisPheLeuLysGln 754
```

RESULT 15

US-11-079-463-7504

```
; Sequence 7504, Application US/11079463
; Publication No. US20060073161A1
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTERIOIDES FRAM
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: PATH00-03DIV2
; CURRENT APPLICATION NUMBER: US/11/079,463
; CURRENT FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/128,705
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: US 09/540,209
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 10444
; SEQ ID NO 7504
; LENGTH: 624
; TYPE: PRT
; ORGANISM: B.fragilis
US-11-079-463-7504
```

```
Alignment Scores:
Pred. No.: 3,85e-10 Length: 624
Score: 180.50 Matches: 57
Percent Similarity: 38.5% Conservative: 35
Best Local Similarity: 23.8% Mismatches: 84
Query Match: 8.4% Indels: 63
DB: 11 Gaps: 6
```

US-10-825-632-4 (1-1197) x US-11-079-463-7504 (1-624)

```
QY 291 TTCTTTTGAAGTACTACTCGATTTCATGTATGGATGCTCTACACCCCTCATGAT 350
Db 356 TrpAsnPheThrAlaSerAspGlyThrGluIleLysGlyMetCysLeuProSer 375
QY 351 CTACACCTCGGAAGAAATATCTACTGTCGTGTTTCATATATGTTGGT----- 398
Db 376 PheAspProAsnLysLysIleProLeuIleValTyrTyrGlyThrProThr 395
QY 398 ----- 398
Db 396 GluArgGlyIleSerAsnProTyrCysAlaGlnLeuPheAlaSerArgAspTyrValVal 415
QY 399 -----CCTCAGGTCAAATA----- 413
Db 416 TyrValIleGlnProSerGlyThrIleGlyPheGlyGlnGluPheSerAlaArgHisVal 435
QY 414 -----GAAATTCAGGATCAGGTGAAGGACTCCAATATATCTAGCTTCT 455
Db 436 AsnAlaTrpGlyLysArgThrAlaAspAspIleIleGluGlyThrLysGlnPheCysLys 455
QY 456 CGATATGATTTCATTAAGTCTAGTCTGTTGGGCATCCACGGCTGCTCTATGGAGGATAC 515
Db 456 GluHisProPheValAsnAspLysLysIleGlyCysLeuGlyAlaSerTyrGlyGlyPhe 475
QY 516 CTCTCCCTGATGATTAATGACAGGTCAGATATCTTCAGGTTGCTATTGCTGGGGCC 575
Db 476 MetThr---GlnTyrLeuGlnThrGlnThrAspIlePheAlaAlaValSerHisAla 494
QY 576 CCAGTCACTCTGGGATCTTCTATGATACAGGATACAGGACGTTATATGGT----- 629
Db 495 GlyIleSerAsnValThrSerTyr-----TrpGlyGluGlyTyrTrpGlyTyrGly 511
QY 630 -----CACCTGACCAAGTGAACAGGCTATTACTACTTA 662
Db 512 TyrAsnAlaIleAlaAlaAspSerTyrProTyrAsnAsnProGluLeuPheThrLys 531
QY 663 GGATCTGTGCCATGACAGCAAGAAAGTTCCCTCTGACCAACCAATCGTTTACTGCTCTTA 722
Db 532 GlnGlySerLeuPheAsnAlaAspLysIleAsnThrPro-----LeuLeuLeuLeu 548
QY 723 CATGTTTCTCTGATGAGATGTCATTTTGCACATACCATATATTTACTGAGTTTTTTTA 782
Db 549 HisGlyThrValAspThrAsnValProIleGlyGluSerIleGlnLeuPheAsnAlaLeu 568
```

Qy 783 GTGAGGCTGGAAGCCATATGATTACAGATCTATCCTCAGGAGACACAGCATA 839
Db 569 LyeIleLeuGlyLysThrValGluPheIleThrValaepGlyGluAsnHisPheIle 587

Search completed: May 2, 2006, 03:10:38
Job time : 37.0056 sec

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.
OM nucleic - nucleic search, using sw model
Run on: May 4, 2006, 02:53:29 ; Search time 8318.8 Seconds
(without alignments)
11404.501 Million cell updates/sec
Title: US-10-825-632-6
Perfect score: 1669
Sequence: 1 aacaggtacagcaaatccta.....aaaaaaaaaaaaaaaa 1669
Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0
Searched: 5883141 seqs, 28421725653 residues
Total number of hits satisfying chosen parameters: 11766282
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Result No.	Score	Query Match	Length	DB ID	Description
1	1669	100.0	1669	6	AR651454
2	1669	100.0	1669	8	AF221636
3	1659	99.4	2668	6	AX405771
4	1659	99.4	2842	6	AX405770
5	1655	99.2	2778	6	AX000290
6	1648.4	98.8	4523	6	AR631281
7	1648.4	98.8	4523	6	AR631281
8	1505	90.2	3125	8	BC030688
9	1501	89.9	4676	6	AR631287
10	1501	89.9	4676	6	AX608743
11	1487.4	89.1	2161	6	BD157001
12	1487.4	89.1	2161	6	AX878058
13	1487.4	89.1	2161	8	AK027826
14	1354	81.1	3127	8	AF221634
15	1353.4	81.1	3143	6	AX354793
16	1349.6	80.9	4829	6	AR631283
17	1349.6	80.9	4829	6	AX608735
18	1347	80.7	3120	6	AR651452

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	DB ID	Description
1	1669	100.0	1669	6	AR651454
2	1669	100.0	1669	8	AF221636
3	1659	99.4	2668	6	AX405771
4	1659	99.4	2842	6	AX405770
5	1655	99.2	2778	6	AX000290
6	1648.4	98.8	4523	6	AR631281
7	1648.4	98.8	4523	6	AR631281
8	1505	90.2	3125	8	BC030688
9	1501	89.9	4676	6	AR631287
10	1501	89.9	4676	6	AX608743
11	1487.4	89.1	2161	6	BD157001
12	1487.4	89.1	2161	6	AX878058
13	1487.4	89.1	2161	8	AK027826
14	1354	81.1	3127	8	AF221634
15	1353.4	81.1	3143	6	AX354793
16	1349.6	80.9	4829	6	AR631283
17	1349.6	80.9	4829	6	AX608735
18	1347	80.7	3120	6	AR651452

19	1347	80.7	4535	8	BC040203
20	1344	80.5	3106	6	AX342633
21	1343.6	80.5	4685	6	AR631288
22	1343.6	80.5	4685	6	AX608745
23	1337	80.1	2797	6	AR448400
24	1250.6	74.9	2510	6	AX338497
25	1234	73.9	3030	8	AY354202
26	1087.4	65.2	2649	6	AX354795
27	1085.8	65.1	2656	8	AY172659
28	1085.8	65.1	2671	6	AR631278
29	1085.8	65.1	2671	6	AX608725
30	924.6	55.4	4799	9	BC043124
31	924.6	55.4	4799	9	BC059222
32	881	52.8	1197	6	AR651453
33	881	52.8	1197	8	AF221635
34	765.2	45.8	4309	6	AR631284
35	765.2	45.8	4309	6	AX608737
36	636.4	38.1	873	6	AX083130
37	605.6	36.3	183341	14	AC161443
38	589.2	35.3	194188	9	AC163387
39	578	34.6	753	6	AX524935
40	524	31.4	735	6	AX524942
41	523.6	31.4	587	6	BD154822
42	523.6	31.4	587	6	AX874760
43	493.4	29.6	823	6	AR263993
44	446.2	26.7	1002	6	AX524939
45	440	26.4	2098	6	CQ721272

ALIGNMENTS

RESULT 1	AR651454	AR651454	1669 bp	DNA	linear	PAT 20-APR-2005
LOCUS	Sequence 6 from patent US 6881564.					
DEFINITION	Sequence 6 from patent US 6881564.					
ACCESSION	AR651454					
VERSION	AR651454.1	GI:62795940				
KEYWORDS	Unknown.					
SOURCE	Unknown.					
ORGANISM	Unclassified.					
REFERENCE	1 (bases 1 to 1669)					
AUTHORS	Abbott, C.A. and Gorrell, M.D.					
TITLE	Dipeptidyl peptidases					
JOURNAL	Patent: US 6881564-A 6 19-APR-2005;					
	The University of Sydney; Sydney;					
FEATURES	AUX;	Location/Qualifiers				
source	1. .1669	/organism="unknown"				
	/mol_type="genomic DNA"					
ORIGIN						
Query Match	100.0%;	Score 1669;	DB 6;	Length 1669;		
Best Local Similarity	100.0%;	Pred. No. 0;				
Matches 1669;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;		
Qy	1	AACAGGTACAGCAATCCTTAAGTCACTTTTAAGATGTCAGAAATAATGATTGATGCTGA	60			
Db	1	AACAGGTACAGCAATCCTTAAGTCACTTTTAAGATGTCAGAAATAATGATTGATGCTGA	60			
Qy	61	AGGAAGGATCATAGATGTCATAGATAAGCACTTAATCAACCTTTTGAGATTTCTATTTGA	120			
Db	61	AGGAAGGATCATAGATGTCATAGATAAGCACTTAATCAACCTTTTGAGATTTCTATTTGA	120			
Qy	121	AGGAGTTGAATATATTCGCCAGAGCTGGACTCTCGAGGAAATAATGCTTGTGTCAT	180			
Db	121	AGGAGTTGAATATATTCGCCAGAGCTGGACTCTCGAGGAAATAATGCTTGTGTCAT	180			
Qy	181	CCTACTAGTCGCTCCAGACTCGCTACAGATAGTGTGATGCTCACCTGAATTTAT	240			
Db	181	CCTACTAGTCGCTCCAGACTCGCTACAGATAGTGTGATGCTCACCTGAATTTAT	240			

Qy	1321	AGAACATTATGAACCTTGCACTCTTTGGCACTACCTTCAAGAAAACCTTGATCAGCTATATGC	1380
Db	1321	AGAACATTATGAACCTTGCACTCTTTGGCACTACCTTCAAGAAAACCTTGATCAGCTATATGC	1380
Qy	1381	TGCTCTAAAAGTGATATAAATTTTGACCTGTGTAGAACTCTCTGTGATACACTTGGCTATTT	1440
Db	1381	TGCTCTAAAAGTGATATAAATTTTGACCTGTGTAGAACTCTCTGTGATACACTTGGCTATTT	1440
Qy	1441	AACCAAAATGAGAGGTTTAAATCAACAGAAAAACACAGAAATTCATCATCATTTTGATAACC	1500
Db	1441	AACCAAAATGAGAGGTTTAAATCAACAGAAAAACACAGAAATTCATCATCATTTTGATAACC	1500
Qy	1501	TGCCATGTAACTACTCTCTGAAATTAATGTGTGTCATGCGAGGGTCTACGGTTTGT	1560
Db	1501	TGCCATGTAACTACTCTCTGAAATTAATGTGTGTCATGCGAGGGTCTACGGTTTGT	1560
Qy	1561	GGTAGTAATCTAATACCTTTAAACCCACATGCTCAAAATCAAAATGATACATATCTCTGAGA	1620
Db	1561	GGTAGTAATCTAATACCTTTAAACCCACATGCTCAAAATCAAAATGATACATATCTCTGAGA	1620
Qy	1621	GACCCAGCAATACCAATAGAAATTACTTAAAAAATAAAAAAATAAAAAA 1669	
Db	1621	GACCCAGCAATACCAATAGAAATTACTTAAAAAATAAAAAAATAAAAAA 1669	
RESULT 2			
AF221636			
LOCUS	AF221636	1669 bp	mRNA linear PRI 05-NOV-2000
DEFINITION	Homo sapiens dipeptidyl peptidase 8 (DPP8) mRNA, partial cds, alternatively spliced.		
ACCESSION	AF221636		
VERSION	AF221636.1 GI:11095191		
KEYWORDS	Homo sapiens (human)		
SOURCE	Homo sapiens		
ORGANISM	Homo sapiens		
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.			
REFERENCE	1 (bases 1 to 1669)		
AUTHORS	Abbott, C.A., Yu, D.M., Woollatt, E., Sutherland, G.R., McCaughan, G.W. and Gorrell, M.D.		
TITLE	Cloning, expression and chromosomal localization of a novel human dipeptidyl peptidase (DPP) IV homolog, DPP8		
JOURNAL	Eur. J. Biochem. 267 (20), 6140-6150 (2000)		
PUBMED	11012666		
REFERENCE	2 (bases 1 to 1669)		
AUTHORS	Abbott, C.A., Yu, D., McCaughan, G.W. and Gorrell, M.D.		
TITLE	Direct Submission		
JOURNAL	Submitted (06-JAN-2000) A.W. Morrow Gastroenterology and Liver Centre, Centenary Institute of Cell Biology and Cancer Medicine, Locked Bag No.6, Newtown, Sydney, NSW 2042, Australia		
FEATURES	Location/Qualifiers		
source	1..1669		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/chromosome="15"		
	/map="15q22"		
	/clone="18"		
	/tissue_type="placenta"		
	/clone_lib="lambda triplexExTM: Clontech"		
gene	<1..1669		
CDS	<1..1399		
	/gene="DPP8"		
	/note="serine protease; member of prolyl oligopeptidase S9b family belonging to enzyme clan SC; isoform missing the catalytic serine required for enzyme activity; additional splice variants of this mRNA are found in GenBank		
	Accession Numbers AF221634, AF221635 and AF221637"		
	/codon_start=2		
	/product="dipeptidyl peptidase 8"		
	/protein_id="AAG29768.1"		

/db_xref="GI:11095192"

/translation="TGTANPKVTFKMSIMIDAEGRIIDVIDKELIQFBEILFEGVEY
IARAGTPEKXWAMSILLDSQRIQVILISPELPIVEDDWMERQRLIESVPSVTP
LIIVEETDIIWINHIDPHVFPQSHHEIEIFASECKTGFRLKYKILSILKESKYKR
SSGLPAPSPDKPIKEEIAITSGWEVLGRHNSIOVDEVRRLVYEGTKDSPLHH
LVVYSVNGEVNLRDRCVSHSCCISOCDPFLSKYSKONKPHCVLSLYKLSPPEDP
TKYKFWANILDSAGLPDYDYPPEIFPFESTGTFTLYGMLYKPHLDQPKKPYTVLF
IYGGPQVATAGAVTLWIFYDTGYTERYHGHDPQNEQGYILGVSVMQAQKFPSPENRL
LLHGFLENDVNHFAHTSILLFLVRAGKPYDLQYPOERHSIRVPSGSEHYELHLHY
LOENLGSRIAALKVI"

ORIGIN

Query Match 100.0%; Score 1669; DB 8; Length 1669;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1669; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	AA	CAGGTACAGCAAACTCTAAAGTCACTTTTAAAGATGTCAGAAATAATGATTGATGCTGA	60
DB	1	AA	CAGGTACAGCAAACTCTAAAGTCACTTTTAAAGATGTCAGAAATAATGATTGATGCTGA	60
QY	61	AG	AAGGATCATAGATGTCATAGATAAGGAACTAATTCAACCTTTTGAGATTCCTATTGGA	120
DB	61	AG	AAGGATCATAGATGTCATAGATAAGGAACTAATTCAACCTTTTGAGATTCCTATTGGA	120
QY	121	AG	GATTTGAATATATTGCGAGAGCTGGATGGAATCTCTGAGGAGAAATAATGCTGGTCCAT	180
DB	121	AG	GATTTGAATATATTGCGAGAGCTGGATGGAATCTCTGAGGAGAAATAATGCTGGTCCAT	180
QY	181	CT	ACTAGATCGCTCCGAGCTCGCTACAGATAGTGTGATCTCACTGCACTGAATATTATAT	240
DB	181	CT	ACTAGATCGCTCCGAGCTCGCTACAGATAGTGTGATCTCACTGCACTGAATATTATAT	240
QY	241	CC	CAGTAAAGATGATGTTATGGAAGGAGAGAGATCAATTGAGTCAGTGCCTGATTCCTGT	300
DB	241	CC	CAGTAAAGATGATGTTATGGAAGGAGAGAGATCAATTGAGTCAGTGCCTGATTCCTGT	300
QY	301	GAG	CCACTAATTTATCTATGAAGAAACAAACAGACATCTGGATTAATATCCATGACATCTT	360
DB	301	GAG	CCACTAATTTATCTATGAAGAAACAAACAGACATCTGGATTAATATCCATGACATCTT	360
QY	361	TC	ATGTTTTTCCCAAGTACAGAGAGAAATGAGTTATTTTGGCTCTGGAATGCAA	420
DB	361	TC	ATGTTTTTCCCAAGTACAGAGAGAAATGAGTTATTTTGGCTCTGGAATGCAA	420
QY	421	AA	CAGGTTCCCGTCAATTTACAAATTTACATCTATTTTAAAGGAAACAAATAATAACG	480
DB	421	AA	CAGGTTCCCGTCAATTTACAAATTTACATCTATTTTAAAGGAAACAAATAATAACG	480
QY	481	AT	CCAGTGTGGTGGCTGCTCCAAAGTATTTCAAGTGTCTATCAAAAGAGGAGATAGC	540
DB	481	AT	CCAGTGTGGTGGCTGCTCCAAAGTATTTCAAGTGTCTATCAAAAGAGGAGATAGC	540
QY	541	AA	TACAGTGTGTAATGGAGAGTTCTTGGCCGGCATGGATCTAATATCCAAAGTTGATGA	600
DB	541	AA	TACAGTGTGTAATGGAGAGTTCTTGGCCGGCATGGATCTAATATCCAAAGTTGATGA	600
QY	601	AG	TACAGAGCTGGTATATTTTGAAGCACCAAGACTCCCTTTTAGAGATCACTCGTA	660
DB	601	AG	TACAGAGCTGGTATATTTTGAAGCACCAAGACTCCCTTTTAGAGATCACTCGTA	660
QY	661	CG	TAGTCAGTTAGCTAAATCCTGGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACA	720
DB	661	CG	TAGTCAGTTAGCTAAATCCTGGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACA	720
QY	721	TT	CTTGCTGATCAGTCAGACTGTGACTTTTAAAGTAAAGTATAGTAACCAAGAGAA	780
DB	721	TT	CTTGCTGATCAGTCAGACTGTGACTTTTAAAGTAAAGTATAGTAACCAAGAGAA	780
QY	781	TC	CACACTGTGCTCCCTTTACAGCTATCAAGTCTCAAGTACCTGACCTGCAAAAC	840
DB	781	TC	CACACTGTGCTCCCTTTACAGCTATCAAGTCTCAAGTACCTGACCTGCAAAAC	840
QY	841	AA	AGGAATTTTGGGCCACCACTTTTGGATTCAGCAGGCTCTCTCTGCTACTATATCTCTCC	900

DB	841	AA	AGGAATTTTGGGCCACCACTTTTGGATTCAGCAGGCTCTCTCTGCTACTATATCTCTCC	900
QY	901	AG	AAATTTTCTCTTTTGAAGTACTACTGATTACATTTGATGGATGCTCTCTACAAGCC	960
DB	901	AG	AAATTTTCTCTTTTGAAGTACTACTGATTACATTTGATGGATGCTCTCTACAAGCC	960
QY	961	TC	ATGATCTACAGCTGGAAGAAATATCTCTGCTGCTTTTATATATATGTTGGTCTCTCA	1020
DB	961	TC	ATGATCTACAGCTGGAAGAAATATCTCTGCTGCTTTTATATATATGTTGGTCTCTCA	1020
QY	1021	GG	TGCTATTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGA	1080
DB	1021	GG	TGCTATTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGA	1080
QY	1081	AC	GTATATGGGTCACTGACCCAGAAATGAACAGGCTATTACTTAGGATCTGTGGCCAT	1140
DB	1081	AC	GTATATGGGTCACTGACCCAGAAATGAACAGGCTATTACTTAGGATCTGTGGCCAT	1140
QY	1141	GC	AAAGCAAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGGTTTCTCTGA	1200
DB	1141	GC	AAAGCAAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGGTTTCTCTGA	1200
QY	1201	TG	AGATGTCATTTTGCACATACCAAGTATATCTGAGTCTTTTAGTGGGCTGGAAA	1260
DB	1201	TG	AGATGTCATTTTGCACATACCAAGTATATCTGAGTCTTTTAGTGGGCTGGAAA	1260
QY	1261	GC	CATATGATTTACAGATCTATCTCAGGAGAGACACAGATAAGATTTCTGAAATCGGG	1320
DB	1261	GC	CATATGATTTACAGATCTATCTCAGGAGAGACACAGATAAGATTTCTGAAATCGGG	1320
QY	1321	AG	AAATATGAACTGTCATCTTTTGCACACTACCTTCAAGAAAACTTTGGATCACTGATTGC	1380
DB	1321	AG	AAATATGAACTGTCATCTTTTGCACACTACCTTCAAGAAAACTTTGGATCACTGATTGC	1380
QY	1381	TG	CTCTAAAAGTGATATAATTTTGAACCTGTGTAGAACTCTCTGCTGATATACACTGGCTATTT	1440
DB	1381	TG	CTCTAAAAGTGATATAATTTTGAACCTGTGTAGAACTCTCTGCTGATATACACTGGCTATTT	1440
QY	1441	AA	CAAAATGAGGAGTTTAAATCAACAGAAAAACAGAAATGATCATCATTTTGTATACC	1500
DB	1441	AA	CAAAATGAGGAGTTTAAATCAACAGAAAAACAGAAATGATCATCATTTTGTATACC	1500
QY	1501	TG	CCATGTAACTACTCTCTGAAAAATAATGTGTGCTGCAATGCAAGGGTCTACGGTTTGT	1560
DB	1501	TG	CCATGTAACTACTCTCTGAAAAATAATGTGTGCTGCAATGCAAGGGTCTACGGTTTGT	1560
QY	1561	GG	TAGTAATCTAATACCTTAAACCCCATGCTCAAAATCAAAATGATATATTTCTCTGAGA	1620
DB	1561	GG	TAGTAATCTAATACCTTAAACCCCATGCTCAAAATCAAAATGATATATTTCTCTGAGA	1620
QY	1621	GACC	CAGCAATACCTAAGAAATTTACTAAAAAATAAAAAAAAAAAAAA 1669	
DB	1621	GACC	CAGCAATACCTAAGAAATTTACTAAAAAATAAAAAAAAAAAAAA 1669	

RESULT 3

LOCUS	AX405771	2668 bp	DNA	linear	PAT 14-JUN-2002
DEFINITION	Sequence 186 from Patent WO0222660.				
ACCESSION	AX405771				
VERSION	AX405771.1	GI:21438981			
KEYWORDS	Homo sapiens (human)				
SOURCE	ORGANISM				
	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;				
	Hominidae; Homo.				
REFERENCE					
AUTHORS	Tang, Y.T., Liu, C., Zhou, P., Asundi, V., Zhang, J., Zhao, Q.A., Ren, F.,				
	Xue, A.T., Yang, Y., Wehrman, T. and Drmanac, R.T.				
TITLE	Novel nucleic acids and polypeptides				
JOURNAL	Patent: WO 0222660-A 186 21-MAR-2002;				

FEATURES		HYSEQ, INC. (US)	
source	Location/Qualifiers		
	1..2668		
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
	234..2408		
CDS	/note="unnamed protein product"		
	/codon_start=1		
	/protein_id="CAD34878.1"		
	/db_xref="GI:21436982"		
	/translation="MAAAMETQLGVIFETADCEENIESQDRPKLEPFVVERYSWSQ LKKLLADRKVGHVYMAKAPHDPMFVKENPDGPHSDRIYVLLAMSENRENTLPFSEI PTXINRAVLMSKPLDLDFQQQLPRLNVLVETSCPNIRMDPKLCPADPDLAFTHSN DTWISNIVTREERRLTYYVHNELANMEDASAGVAPVLQDEEDRYSGVMCPKAEIT PGGKILRIYEENDESEVELIHVTSPLMTERADSPRYPKTGTPANPKPTFKMSIMI DAGRIILVIDKELIOPREILFEQVEVILARAGTTPBKVYWSILLDRSOTRLQIVLIS PLIFPVEDDUMERORLIESVPDSVTBLIYEETDMINIHDIHVPVROSHERIEF IFASCKTGFRHLXYLTSILKESRYKSSGGLPAPSDPKPIKEETATISGEWEVLGR HGSNIOQVEVRLVTFEGTKOSPLEHLLTVSYNVFGEVTRLDRGYSKSCISQHCD FISIKSNQNFPHDLQCKKYPFVLFIYGGPQVATAGAEVILWIFDIYGYTERHGH TFGTLYGMLKPHDLQCKKYPFVLFIYGGPQVATAGAEVILWIFDIYGYTERHGH PQNSQGYTLGSVMAQBEKFPSEPNRLLLHGLDENVHFATSIILLSPFLVRAGRPYD LQIYPOERHSIRVPESGEHYELHLLHQLNELGSRIAALKVI"		
ORIGIN			
Query Match		99.4%; Score 1659; DB 6; Length 2668;	
Best Local Similarity		100.0%; Pred. No. 0;	
Matches 1659; Conservative		0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1 AACAGGTACAGCAATTCCTAAAGTCACCTTTTAAAGATGTGAGAAATATGATGATGCTGA	60	
Db	1010 AACAGGTACAGCAATTCCTAAAGTCACCTTTTAAAGATGTGAGAAATATGATGATGCTGA	1069	
Qy	61 AGGAAGGATCATAGATGTCATAGATAAGGAACATAATCAACCTTTTTCAGAGATTCATTTGA	120	
Db	1070 AGGAAGGATCATAGATGTCATAGATAAGGAACATAATCAACCTTTTTCAGAGATTCATTTGA	1129	
Qy	121 AGGAGTTGAATATATTGCGAGAGCTGGATGGACTCTCGAGGGAATAATGCTTGGTCCAT	180	
Db	1130 AGGAGTTGAATATATTGCGAGAGCTGGATGGACTCTCGAGGGAATAATGCTTGGTCCAT	1189	
Qy	181 CCTACTAGATCGCTCCAGACTCGCTACAGATAGTTGATCTCACCTGAATATTATAT	240	
Db	1190 CCTACTAGATCGCTCCAGACTCGCTACAGATAGTTGATCTCACCTGAATATTATAT	1249	
Qy	241 CCCAGTAGAGATGATGTTATGGAAGCGCAGAGCTCATTTGAGTCAGTGCCTGATCTGT	300	
Db	1250 CCCAGTAGAGATGATGTTATGGAAGCGCAGAGCTCATTTGAGTCAGTGCCTGATCTGT	1309	
Qy	301 GAGCCACTAATTTATCTATGAAGAACACAGACATCTGGATAAATATCCATGACATCTT	360	
Db	1310 GAGCCACTAATTTATCTATGAAGAACACAGACATCTGGATAAATATCCATGACATCTT	1369	
Qy	361 TCATGTTTTTCCCAAGTCACGAGAGGAAATTTGAGTTATTTTTCCTCTGGAATGCAA	420	
Db	1370 TCATGTTTTTCCCAAGTCACGAGAGGAAATTTGAGTTATTTTTCCTCTGGAATGCAA	1429	
Qy	421 AACAGGTTCCCGTCATTTTACAAAATTTACATCTATTATTTAAAGGAAAGCAAAATATAACG	480	
Db	1430 AACAGGTTCCCGTCATTTTACAAAATTTACATCTATTATTTAAAGGAAAGCAAAATATAACG	1489	
Qy	481 ATCCAGTGTGGGCTCCCTGCTCCCAAGTGATTTCAAGTGTCTCTATCAAGAGGAGATAGC	540	
Db	1490 ATCCAGTGTGGGCTCCCTGCTCCCAAGTGATTTCAAGTGTCTCTATCAAGAGGAGATAGC	1549	
Qy	541 AATTACAGTGTGGAATGGGAAGTTCTTGGCCGGCATGGATCTAATATCCAAAGTTGATGA	600	
Db	1550 AATTACAGTGTGGAATGGGAAGTTCTTGGCCGGCATGGATCTAATATCCAAAGTTGATGA	1609	
Qy	601 AGTCAGAGGCTGGTATATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACCTGTA	660	
Db	1610 AGTCAGAGGCTGGTATATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACCTGTA	1669	

RESULT 4
AX405770
LOCUS

AX405770 2842 bp DNA linear PAT 14-JUN-2002

DEFINITION	Sequence 185 from Patent WO0222660.	
ACCESSION	AX405770	
VERSION	AX405770.1	GI:21438979
KEYWORDS		
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	
REFERENCE		
AUTHORS	Tang, Y. T., Liu, C., Zhou, P., Asundi, V., Zhang, J., Zhao, Q. A., Ren, F., Xue, A. J., Yang, Y., Wehrman, T. and Drmanac, R. T.	
TITLE	Novel nucleic acids and polypeptides	
JOURNAL	Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.	
FEATURES		
source	1..2842	
	/organism="Homo sapiens"	
	/mol_type="unassigned DNA"	
	/db_xref="taxon:9606"	
CDS	234..2582	
	/note="unnamed protein product"	
	/codon_start=1	
	/protein_id="CAD34877.1"	
	/db_xref="GI:21438980"	
	/translations="MAAAMETBQLGVEIFETADCEENIESQDRPKLBPFFVYRWSQ LKLLADTRRYHGMKAPHDIFMFKRNDPDGPHSDRIYYLAMSGNRENTLFYSEI PKTINRAALVHWKPLDLFOATLDYGMYSREELLRRKRI GTVGIASVDYHQSG TFLPQAGSGLVHWDGSGPGQFTQPLRPNLVETSCPNRMDPKLCPADPDWIAFIHNS DIWLSNIVTEERRLTYVHNLNMEEDARSAGVATFVLQEFPRYSGYWCAPKAEIT PSGKILIRLIEBDESEVEI IHVTSFMLETRRDSFRYPKGTANPKVTFPKWSEIMI DAEGRIDVDEKLIQPFILFEGVEYIARAGWTEPGKYAWSIILDRSQRLOVLIS PELFIPEVDVMERORLIESVPSVPLIIYEETDIDINIHDIFHVFPQSHBEEIEF IPASECKTGRHLVYKITSILKESKYKRSGGLPAPSPDKFCPIKEEIAITSGEWELGR HGSNIQVDEVRRLVPEGTQDPLHLVYVSNVNGEVETRLDRGYSHSCCI SOHCD FFTSKYNQKNPCIVSLYKLSPEDDPTCKTFEFWATILDSAGLPUPPEIPFESFS TGTTGLYKPHDLQPKKYPTVLFYIGGQVAIAGAPVLMIFDYDTGYTERYMGH PDQNEQYILGYSVMAQKFPSEPNRLHLLHGLFDENVHFAHTSILLSFLVRAGKPYD LQIYQPRHSIRVPESGEHYELHLLHYLQENLGSRIAALKVI"	
ORIGIN		
Query Match	99.4%; Score 1659; DB 6; Length 2842;	
Best Local Similarity	100.0%; Pred. No. 0;	
Matches 1659; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 AACAGGTACAGCAATCCTAAAGTCACCTTTTAAGATGTCAGAAATAATGATTGATCTGA	60
DB	1184 AACAGGTACAGCAATCCTAAAGTCACCTTTTAAGATGTCAGAAATAATGATTGATCTGA	1243
QY	61 AGGAAGGATCATAGATGTCATAGATAAGGAACCTTAATTCACCTTTTGAGATTCATTGGA	120
DB	1244 AGGAAGGATCATAGATGTCATAGATAAGGAACCTTAATTCACCTTTTGAGATTCATTGGA	1303
QY	121 AGGAGTTGAATATATCCAGAGCTGATGGACTCTCAGAGGAAATAATGCTGGTCCAT	180
DB	1304 AGGAGTTGAATATATCCAGAGCTGATGGACTCTCAGAGGAAATAATGCTGGTCCAT	1363
QY	181 CCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGGAATATTAT	240
DB	1364 CCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGGAATATTAT	1423
QY	241 CCCAGTGAAGATGATGTTATGGAAAGGACAGACTCATTAGTCAAGTGCCTGATCTCTGT	300
DB	1424 CCCAGTGAAGATGATGTTATGGAAAGGACAGACTCATTAGTCAAGTGCCTGATCTCTGT	1483
QY	301 GACGCCACTAATTATCTATGAAGAAACACAGACATCTGGATAAATATCCATGACATCTT	360
DB	1484 GACGCCACTAATTATCTATGAAGAAACACAGACATCTGGATAAATATCCATGACATCTT	1543
QY	361 TCATGTTTTTCCCCAAGTCAACGAGGAAATGAGTTATTTTTCCTCTGATGCA	420
DB	1544 TCATGTTTTTCCCCAAGTCAACGAGGAAATGAGTTATTTTTCCTCTGATGCA	1603

Db 2132 GGTGCTATGCTGGGGCCCCAGTCACCTCTGTGGATCTTCTATGATACAGGATACACGGA 2191
Qy 1081 ACCTTATATGGGTCAACCTTGACAGAAATGAACAGGGCTATTACTTAGGATCTGTGGCCAT 1140
Db 2192 ACCTTATATGGGTCAACCTTGACAGAAATGAACAGGGCTATTACTTAGGATCTGTGGCCAT 2251
Qy 1141 GCAGCAGAAAATGTTCCCTCTGAACCAAAATCGTTTACTGCTCTTACATGGTTTCTCGGA 1200
Db 2252 GCAGCAGAAAATGTTCCCTCTGAACCAAAATCGTTTACTGCTCTTACATGGTTTCTCGGA 2311
Qy 1201 TGAGATGTCCTATTTTGACATACACAGTATATTACTGAGTTTCTTACTGAGGGCTCGAAA 1260
Db 2312 TGAGATGTCCTATTTTGACATACACAGTATATTACTGAGTTTCTTACTGAGGGCTCGAAA 2371
Qy 1261 GCATATGATTTTACAGATCTATCTCTCAGGAGAGACACAGATAGAGTTCTCTGAATCGGG 1320
Db 2372 GCATATGATTTTACAGATCTATCTCTCAGGAGAGACACAGATAGAGTTCTCTGAATCGGG 2431
Qy 1321 AGAACATTAATGAATGCTATCTTTTGCACTACCTTCAAGAAAACTTTGGATCAGTATGTC 1380
Db 2432 AGAACATTAATGAATGCTATCTTTTGCACTACCTTCAAGAAAACTTTGGATCAGTATGTC 2491
Qy 1381 TGCCTTAAAGTGATATAATTTTGAACCTGCTGAGACTCTCTGGTATACACTGGCTATT 1440
Db 2492 TGCCTTAAAGTGATATAATTTTGAACCTGCTGAGACTCTCTGGTATACACTGGCTATT 2551
Qy 1441 AACCAATGAGGAGGTTTAAATCAACAGAAAAACACAGAAATGATCATCACATTTTGATACC 1500
Db 2552 AACCAATGAGGAGGTTTAAATCAACAGAAAAACACAGAAATGATCATCACATTTTGATACC 2611
Qy 1501 TGCCATGTAACATCTACTCTGAAATTAATGTTGGTGCCATGCAAGGGTCTACGGTTTGT 1560
Db 2612 TGCCATGTAACATCTACTCTGAAATTAATGTTGGTGCCATGCAAGGGTCTACGGTTTGT 2671
Qy 1561 GGTAGTAATCTAATACCTTTAACCACCATGCTCAAAATCAATGATACATATTCCTGAGA 1620
Db 2672 GGTAGTAATCTAATACCTTTAACCACCATGCTCAAAATCAATGATACATATTCCTGAGA 2731
Qy 1621 GACCCAGCAATACCAATAGAAATTTACTAAAAAATTTTAAAAAATTTTAAAAA 1667
Db 2732 GACCCAGCAATACCAATAGAAATTTACTAAAAAATTTTAAAAAATTTTAAAAA 2778

RESULT 6
AR631281
LOCUS AR631281 4523 bp DNA linear PAT 14-FEB-2005
DEFINITION Sequence 8 from patent US 6844180.
ACCESSION AR631281
VERSION AR631281.1 GI:59770924
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 4523)
AUTHORS Oi,S., Akinsanya,K.O., Riviere,P.J.M. and Junien,J.-L.
TITLE Serine protease genes related to DPPIV
JOURNAL Patent: US 6844180-A 8 18-JAN-2005;
Ferring BV;
NLX;
FEATURES
source Location/Qualifiers
1. .4523
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 98.8%; Score 1648.4; DB 6; Length 4523;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1663; Conservative 0; Mismatches 1; Indels 4; Gaps 1;
Qy 1 AACAGGTACAGCAATCTTAAGTCACTTTTAAAGTGTACAGAAATATGATGCTGA 60
Db 1164 AACAGGTACAGCAATCTTAAGTCACTTTTAAAGTGTACAGAAATATGATGCTGA 1223

Qy 61 AGAAGGATCATAGATGTCATAGATAAGAACTAATTCACCTTTGAGATCTTATTGTA 120
Db 1224 AGAAGGATCATAGATGTCATAGATAAGAACTAATTCACCTTTGAGATCTTATTGTA 1283
Qy 121 AGGAGTTGAATATATTGCCAGAGCTGGATGGACTCCTCGAGGAAAAATATGCTTGGTCCAT 180
Db 1284 AGGAGTTGAATATATTGCCAGAGCTGGATGGACTCCTCGAGGAAAAATATGCTTGGTCCAT 1343
Qy 181 CCTACTAGATCGCTCCAGACTCGGCTACAGATAGTGTGATCTCACCTGAAATATTAT 240
Db 1344 CCTACTAGATCGCTCCAGACTCGGCTACAGATAGTGTGATCTCACCTGAAATATTAT 1403
Qy 241 CCCAGTAGAAGATGATGTTATGGAAGGACAGAGACTCAATGAGTCAGTGCCTGATCTGT 300
Db 1404 CCCAGTAGAAGATGATGTTATGGAAGGACAGAGACTCAATGAGTCAGTGCCTGATCTGT 1463
Qy 301 GAGCCACTAATTTATCTATGAAGAAACACAGACATCTGGATAAATATCCATGACATCTT 360
Db 1464 GAGCCACTAATTTATCTATGAAGAAACACAGACATCTGGATAAATATCCATGACATCTT 1523
Qy 361 TCATGTTTTTCCCAAGTCACGAAGAGGAAATTCAGTTTATTTTGGCTCTGAATGCAA 420
Db 1524 TCATGTTTTTCCCAAGTCACGAAGAGGAAATTCAGTTTATTTTGGCTCTGAATGCAA 1583
Qy 421 AACAGTTTCCGTCATTTATACAAAATTTACATCTATTTTAAAGGAAACAAATATAACG 480
Db 1584 AACAGTTTCCGTCATTTATACAAAATTTACATCTATTTTAAAGGAAACAAATATAACG 1643
Qy 481 ATCCAGTGGTGGCTGCCCTGCCAAGTGATTTCAAGTGTCTCTATCAAGAGGAGATAGC 540
Db 1644 ATCCAGTGGTGGCTGCCCTGCCAAGTGATTTCAAGTGTCTCTATCAAGAGGAGATAGC 1703
Qy 541 AATTACAGTGGTGAATGGGAAAGTTCTTGGCCGGCATGGATCTAATATCCAAAGTTGATGA 600
Db 1704 AATTACAGTGGTGAATGGGAAAGTTCTTGGCCGGCATGGATCTAATATCCAAAGTTGATGA 1763
Qy 601 AGTCAGAGGCTGGTATATTTTGAAGGCAACAAAGCTCCCTTTAGAGCATCACTGTGA 660
Db 1764 AGTCAGAGGCTGGTATATTTTGAAGGCAACAAAGCTCCCTTTAGAGCATCACTGTGA 1823
Qy 661 CGTAGTCAGTTACGTAAATCCCTGGAGAGGTGACAAGGCTGACTGACCTGGCTACTCACA 720
Db 1824 CGTAGTCAGTTACGTAAATCCCTGGAGAGGTGACAAGGCTGACTGACCTGGCTACTCACA 1883
Qy 721 TTCTTGTCATCAGTCAGCACTGTGACTTCTTTTAAAGTAAAGTATAGTATACCAAGAGAA 780
Db 1884 TTCTTGTCATCAGTCAGCACTGTGACTTCTTTTAAAGTAAAGTATAGTATACCAAGAGAA 1943
Qy 781 TCCACACTGTGTCTCCCTTTAAGCTATCAAGCTCCTGAAGATGACCCAACTTGCAGAAC 840
Db 1944 TCCACACTGTGTCTCCCTTTAAGCTATCAAGCTCCTGAAGATGACCCAACTTGCAGAAC 2003
Qy 841 AAAGGAATTTGGGCCACCATTTTGGATTCAGCAGGTCCTCTTCTGACTATATCTCTCC 900
Db 2004 AAAGGAATTTGGGCCACCATTTTGGATTCAGCAGGTCCTCTTCTGACTATATCTCTCC 2063
Qy 901 AGAAATTTCTCTTTTGAAGTACTACTGATTTTACATTTGATGGGATGCTCTCAAGGCC 960
Db 2064 AGAAATTTCTCTTTTGAAGTACTACTGATTTTACATTTGATGGGATGCTCTCAAGGCC 2123
Qy 961 TCATGATCTACAGCCTGGAAAGAAATATCTACTGTGTGTTTCATATATGTTGGTCTCTCA 1020
Db 2124 TCATGATCTACAGCCTGGAAAGAAATATCTACTGTGTGTTTCATATATGTTGGTCTCTCA 2179
Qy 1021 GGTGCTATTGCTGGGGCCCCAGTCACTCTGTGGATCTCTTATGATACAGGATACAGGA 1080
Db 2180 GGTGCTATTGCTGGGGCCCCAGTCACTCTGTGGATCTCTTATGATACAGGATACAGGA 2239
Qy 1081 ACCTTATATGGGTCAACCTTGACAGAAATGAACAGGGCTATTACTTAGGATCTGTGGCCAT 1140
Db 2240 ACCTTATATGGGTCAACCTTGACAGAAATGAACAGGGCTATTACTTAGGATCTGTGGCCAT 2299
Qy 1141 GCAAGCAGAAAAGTTCCCTCTGAAACCAATCGTTTACTGTCTTTACATGTTTCTCTGGA 1200

```
Db 2300 GCAAGCAGAAAGTTCCTCCCTCTGACCAAAATCGTTTACTGCTCTTACATGCTTCTCTGGA 2359
Qy 1201 TGAGATGTCCTATTTGGACATACAGATATATTAAGTATGTTTATGAGGCTGGAA 1260
Db 2360 TGAGATGTCCTATTTGGACATACAGATATATTAAGTATGTTTATGAGGCTGGAA 2419
Qy 1261 GCCATATGATTTACAGATCTATCCTCAGGAGAGACACAGCATTAAGATGTTCTGAATCGGG 1320
Db 2420 GCCATATGATTTACAGATCTATCCTCAGGAGAGACACAGCATTAAGATGTTCTGAATCGGG 2479
Qy 1321 AGAACATTAATGAACATGATCTTTTGGACATACCTTCAAGAAAACTTGGATCAAGTATGC 1380
Db 2480 AGAACATTAATGAACATGATCTTTTGGACATACCTTCAAGAAAACTTGGATCAAGTATGC 2539
Qy 1381 TGCTCTAAAAGTGATATAAATTTTGACCTGTGTAGAACTCTCTGGTATACATCGCTATTT 1440
Db 2540 TGCTCTAAAAGTGATATAAATTTTGACCTGTGTAGAACTCTCTGGTATACATCGCTATTT 2599
Qy 1441 AACCAAATGAGGAGGTTTAATCAACAGAAAAACAGAAATTGATCATCACATTTTGATACC 1500
Db 2600 AACCAAATGAGGAGGTTTAATCAACAGAAAAACAGAAATTGATCATCACATTTTGATACC 2659
Qy 1501 TGCCATGTAACTACTCTCTGAAATAAATGTGTGCTGCAATGCAAGGGGTCTACGGTTTGT 1560
Db 2660 TGCCATGTAACTACTCTCTGAAATAAATGTGTGCTGCAATGCAAGGGGTCTACGGTTTGT 2719
Qy 1561 GGATGTAATCTTAATACCTTAAACCCACATGCTCAAAATCAAAATGATACATATTCCTGAGA 1620
Db 2720 GGATGTAATCTTAATACCTTAAACCCACATGCTCAAAATCAAAATGATACATATTCCTGAGA 2779
Qy 1621 GACCCAGCAATACCATAGATTACTAAABAAAAAABAAAAAABAAAAA 1668
Db 2780 GACCCAGCAATACCATAGATTACTAAABAAAAAABAAAAAABAAAAA 2827

RESULT 7
AX608731
LOCUS AX608731 4523 bp DNA linear PAT 17-FEB-2003
DEFINITION Sequence 8 from Patent WO0231134.
ACCESSION AX608731
VERSION AX608731.1 GI:28404301
KEYWORDS Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
1
Q1.S., Akinsanya.K.O., Riviere,P.J. and Junien,J.L.
Novel serine protease genes related to dppiv
Patent: WO 0231134-A 8 18-APR-2002;
Ferring BV (NL)
FEATURES
    source
        1..4523
            /organism="Homo sapiens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9606"

ORIGIN
Query Match 98.8%; Score 1648.4; DB 6; Length 4523;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1663; Conservative 1; Mismatches 4; Gaps 1;

Qy 1 AACAGTACAGCAATCCTTAAGTCACCTTTTAAAGATGTCAGAAATTAATGATTGATGCTGA 60
Db 1164 AACAGTACAGCAATCCTTAAGTCACCTTTTAAAGATGTCAGAAATTAATGATTGATGCTGA 1223
Qy 61 AGGAGGATCATAGATGTCATAGATAAGGACATTAATTCACCTTTTGAAGATCTATTGGA 120
Db 1224 AGGAGGATCATAGATGTCATAGATAAGGACATTAATTCACCTTTTGAAGATCTATTGGA 1283
Qy 121 AGGAGTGAATATATTGCCAGAGCTGGATGCCCTCCTGAGGAAAAATATGCTTGGTCCAT 180
```

```
Db 1284 AGAGTTGAATATATTTGCCAGAGCTGGATGGACTCTGAGGAAAAATATGCTTGGTCCAT 1343
Qy 181 CCTACTAGATCGCTCCAGACTCGCCTACAGATAGTGTGATCTCACCTGAATATTATTTAT 240
Db 1344 CCTACTAGATCGCTCCAGACTCGCCTACAGATAGTGTGATCTCACCTGAATATTATTTAT 1403
Qy 241 CCCAGTAGAAGATGATGTTTATGAAAGGCGAGAGACTCATTTGAGTCAGTGCCTGATCTGT 300
Db 1404 CCCAGTAGAAGATGATGTTTATGAAAGGCGAGAGACTCATTTGAGTCAGTGCCTGATCTGT 1463
Qy 301 GAGCCCACTAATTTATCTATGAAAGAAACAAACAGACATCTGGATAAATATCCATGACATCTT 360
Db 1464 GAGCCCACTAATTTATCTATGAAAGAAACAAACAGACATCTGGATAAATATCCATGACATCTT 1523
Qy 361 TCATGTTTTTCCCAAGTTCACGAAAGAGGAAATTCAGTATTTATTTTGGCTCTGAATGCAA 420
Db 1524 TCATGTTTTTCCCAAGTTCACGAAAGAGGAAATTCAGTATTTATTTTGGCTCTGAATGCAA 1583
Qy 421 AACAGGTTTCCCGTCATTTATACAAAATTTACATCTATTTTAAAGGAAAGCAAAATATAAACG 480
Db 1584 AACAGGTTTCCCGTCATTTATACAAAATTTACATCTATTTTAAAGGAAAGCAAAATATAAACG 1643
Qy 481 ATCCAGTGGTGGGCTCGCTGCTCCAAAGTGATTTCAAGTGTCTTATCAAGAGGAGATAGC 540
Db 1644 ATCCAGTGGTGGGCTCGCTGCTCCAAAGTGATTTCAAGTGTCTTATCAAGAGGAGATAGC 1703
Qy 541 AATTACCAGTGGTGAATGGAAGTTCTTGGCGCGCATGATCTAATATCCAAAGTTGATGA 600
Db 1704 AATTACCAGTGGTGAATGGAAGTTCTTGGCGCGCATGATCTAATATCCAAAGTTGATGA 1763
Qy 601 AGTCAGAGGCTGGTATATTTTGAAGGACACAAAGACTCCCTTTAGAGCATACCTGTGA 660
Db 1764 AGTCAGAGGCTGGTATATTTTGAAGGACACAAAGACTCCCTTTAGAGCATACCTGTGA 1823
Qy 661 CGTAGTCAGTACGTTAAATCTCTGGAGAGGTGACAAAGGCTGACTGACCGTGGCTACTCACA 720
Db 1824 CGTAGTCAGTACGTTAAATCTCTGGAGAGGTGACAAAGGCTGACTGACCGTGGCTACTCACA 1883
Qy 721 TTCTGCTGATCAGTCAGCAGCTGTGACTCTTTTAAAGTAAGTATAGTACCAAGAGAA 780
Db 1884 TTCTGCTGATCAGTCAGCAGCTGTGACTCTTTTAAAGTAAGTATAGTACCAAGAGAA 1943
Qy 781 TCACACTGTGTGCTCCTTTTCAAGCTATCAAGTCTGGAAGATGACCCAACTTGTGCAAAAC 840
Db 1944 TCACACTGTGTGCTCCTTTTCAAGCTATCAAGTCTGGAAGATGACCCAACTTGTGCAAAAC 2003
Qy 841 AAAGGAATTTTGGGCCACCAATTTTGGATTTCAAGAGGCTCTCTCTGACTATATCTCTCC 900
Db 2004 AAAGGAATTTTGGGCCACCAATTTTGGATTTCAAGAGGCTCTCTCTGACTATATCTCTCC 2063
Qy 901 AGAAATTTTCTCTTTTGAAGTACTACTGGAATTTTACATTTGATGGGATGCTCTACAAAGCC 960
Db 2064 AGAAATTTTCTCTTTTGAAGTACTACTGGAATTTTACATTTGATGGGATGCTCTACAAAGCC 2123
Qy 961 TCATGATCTACAGCCTGGAAAGAAATATCTTACTGTGCTGTTCATATATGTTGGTCTCTCA 1020
Db 2124 TCATGATCTACAGCCTGGAAAGAAATATCTTACTGTGCTGTTCATATATGTTGGTCTCTCA 2179
Qy 1021 GGTGTCTATTGCTGGGGCCCCAGTCACCTCTGTGATCTTCTATGATACAGGATACACGGA 1080
Db 2180 GGTGTCTATTGCTGGGGCCCCAGTCACCTCTGTGATCTTCTATGATACAGGATACACGGA 2239
Qy 1081 ACCTTATATGGTTCACCTTCAGCAAGTGAACAGGGCTATTACTTAGGATCTGTGGCCAT 1140
Db 2240 ACCTTATATGGTTCACCTTCAGCAAGTGAACAGGGCTATTACTTAGGATCTGTGGCCAT 2299
Qy 1141 GCAGCAGAAATGTTCCCTCTGACCAAAATCGTTTACTGCTCTTACTGATCTGTCTGGA 1200
Db 2300 GCAGCAGAAATGTTCCCTCTGACCAAAATCGTTTACTGCTCTTACTGATCTGTCTGGA 2359
Qy 1201 TGAGATGTCATTTTGGACATACAGGTATATTAATGATGTTTTTTAGTGGGCTGAAA 1260
```

Db	2360	TGGAATGTCCTATTTGGACATACACAGTATATTACTAGATTTTTTTAGTGAGGGCTGAAA	2419
Qy	1261	GCCATATGATTTACAGATCTATCTCAGGAGAGACACAGCATAGAGTTCTGATCGGG	1320
Db	2420	GCCATATGATTTACAGATCTATCTCAGGAGAGACACAGCATAGAGTTCTGATCGGG	2479
Qy	1321	AGAACATTAAGTACATCTTTTGGCACTACCTTCAAGAAAACCTTGGATCAGTATTC	1380
Db	2480	AGAACATTAAGTACATCTTTTGGCACTACCTTCAAGAAAACCTTGGATCAGTATTC	2539
Qy	1381	TGCTCTAAAAGTGATATAATTTTGACCTGTGTGAGAACTCTCTGGTATACACTGGCTATTT	1440
Db	2540	TGCTCTAAAAGTGATATAATTTTGACCTGTGTGAGAACTCTCTGGTATACACTGGCTATTT	2599
Qy	1441	AACCAATAGGAGGTTTAAATCAACAGAAAACACAGAATTGATCATCATCTTTTGATACC	1500
Db	2600	AACCAATAGGAGGTTTAAATCAACAGAAAACACAGAATTGATCATCATCTTTTGATACC	2659
Qy	1501	TGCATGTACATCTACTCTGAAATAAATGTTGGTGCATGCAGGGGTCTACGGTTGT	1560
Db	2660	TGCATGTACATCTACTCTGAAATAAATGTTGGTGCATGCAGGGGTCTACGGTTGT	2719
Qy	1561	GGTAGTAATCTAATACCTTAAACCCACATGCTCAAAATCAAAATGATACATATTCCTGAGA	1620
Db	2720	GGTAGTAATCTAATACCTTAAACCCACATGCTCAAAATCAAAATGATACATATTCCTGAGA	2779
Qy	1621	GACCCAGCAATACCATAGAAATTTACTAAAAAATAAAAAAAAAAAAAAAAAAAAAA	1668
Db	2780	GACCCAGCAATACCATAGAAATTTACTAAAAAATAAAAAAAAAAAAAAAAAAAAAA	2827
RESULT 8			
BC030688			
LOCUS			
DEFINITION	BC030688	3125 bp mRNA linear PRI 28-JUL-2005	
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
1 (bases 1 to 3125)			
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,			
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,			
Altschul, S.F., Ziesberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,			
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,			
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, D.,			
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,			
Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,			
Carninci, P., Prange, C., Raha, S., Loquellano, N.A., Peters, G.J.,			
Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J.,			
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,			
Worley, K.C., Hale, S., Garcia, A.M., Gay, D.J., Hulyk, S.W.,			
Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,			
Fahy, J., Helton, E., Kettman, M., Madan, A., Young, A.C., Shevchenko, Y.,			
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,			
Bouffard, G.O., Blakesley, R.W., Touchman, J.W., Green, E.D.,			
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,			
Butterfield, Y.S., Krywinski, M.I., Skalska, U., Skalska, D.E.,			
Schneerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.			
Mammalian Gene Collection Program Team			
Generation and initial analysis of more than 15,000 full-length			
human and mouse cDNA sequences			
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)			
12477932			
2 (bases 1 to 3125)			
CONSTRM			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
CONSTRM			
TITLE			
JOURNAL			

REMARK			
COMMENT			
Gene Collection (MGC), Bethesda, MD 20892-2590, USA			
NIH-MGC Project URL: http://mgc.ncl.nih.gov			
On Aug 25, 2003 this sequence version replaced gi:21265132.			
Contact: MGC help desk			
Email: cgabs-x@mail.nih.gov			
Tissue Procurement: Niklos Palkovits, M.D., Ph.D.			
cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki			
Toshiyuki and Piero Carninci (RIKEN)			
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)			
DNA Sequencing by: Institute for Systems Biology			
http://www.systemsbio.org			
contact: amad@systemsbiology.org			
Anup Madan, Jessica Fahey, Erin Helton, Mark Kettman, Anuradha			
Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting			
Clone distribution: MGC clone distribution information can be found			
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov			
Series: IRAC Plate: 33 Row: d Column: 5			
This clone was selected for full length sequencing because it			
passed the following selection criteria: matched mRNA gi: 37577092.			
Location/Qualifiers			
1. 3125			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/clone="MGC:26191 IMAGE:4822550"			
/tissue_type="Testis"			
/clone_lib="NIH MGC 97"			
/lab_host="DH10B"			
/note="Vector: pBluescriptR"			
1. 3125			
/gene="DPP8"			
/notes="synonyms: FLJ20283, FLJ14920, MGC26191, MSTP141"			
/db_xref="GeneID:54878"			
/db_xref="MIM:606819"			
319..2862			
/gene="DPP8"			
/codon_start=1			
/product="dipeptidyl peptidase 8, isoform 4"			
/protein_id="AAH30688.3"			
/db_xref="GI:54114905"			
/db_xref="GeneID:54878"			
/db_xref="MIM:606819"			
/translation="MWKSEQMKIKSGKNMAAAWETQLGVETPETADCENIESQD			
RPKLEPFYERYSWSQLKLADTKYGMMAKAPHPDMFVKNDPDPHSRIYLL			
AMGSENNRTLFYSEIPTKINRAAAILMSKPLDLFOATLDYGYSEELRERKR			
IGTVGIASYDYHQSGTFLFOAGSGIYHVKGQPGCTQQLRPNLVETSCPNRMDP			
KLPADPDIAFIHSDNIWISNIVFREERLLTYVHNELANNEEDARSAGVATFVLOEE			
FDYSGYWCPCRAETTPSGGKILRIYEENDESEVEIIVTSPMLETRADSPRYPKT			
GTANPKVTFMSEIMIDAGRIIDVIDKLIQPFELFEGVEYIARAGWTPEGKAW			
ILLDRSQTRLQIVLISPELFIPEDDVMERQLITESVPSVPLIIEETTDWINIH			
DIHVPFQSHHEBIEIFASECKTGRHLKYITSLKESKYRSGGLPAPSPDKCPI			
KEBIAITSGEWEVLGRHGSNIQVDEVRRLVPEGTQDSPLSHLYVSVYVNPGBVRL			
TDYGSCHSCISQHCDFITSKSNOKNPHCVSLYKSSPEDDPTCKTKEFWATLDSA			
GVLPDTPPEISFESTGTFTYNGMLYKPHDLQPKCKPTVILFYGGPQVQLVNNREK			
GLPYFRNLTLASLGIVVVVDNRKSGHRLKPEGAFKYMVAIAGAPVILWIFDTGY			
TERYMGPPQNEQGYLGVSVAQAEKFPSEPNRLLLHFLDENVFAHTSILLSLFLV			
RAGKPYDLQIYQPERHSIRVPSGSEHYELHLLHYLQENLGSRIAALKVI"			
ORIGIN			
Query Match	90.2%	Score 1505;	DB 8; Length 3125;
Best Local Similarity	91.9%	Pred. No. 0;	
Matches 1662; Conservative	0;	Mismatches	0; Indels 147; Gaps 1;
Qy	1	AACAGGTACAGCAAACTCTAAAGTCACCTTTTAAGATGTCAGAAATAATGATGCTGA	60
Db	1317	AACAGGTACAGCAAACTCTAAAGTCACCTTTTAAGATGTCAGAAATAATGATGCTGA	1376
Qy	61	AGGAAGGATCATAGATGTCATAGATAAGGAACCTTAATCAACCTTTTGAGATTCATTGA	120
Db	1377	AGGAAGGATCATAGATGTCATAGATAAGGAACCTTAATCAACCTTTTGAGATTCATTGA	1436
Qy	121	AGGAGTTGAATATATTGCCAGAGCTGGATGGACTCTCTCTGGGAAAAATATGCTGTGTCAT	180

```
Db 1437 AGAGTTGGAATATATTGCCAGAGCTGGATGGACTCCTGAGGGAATAATGCTTGGTCCAT 1496
Qy 181 CCTACTAGATCGCTCCAGACTCGCTTACAGATAGTGTGTACCTCACTGAATATTAT 240
Db 1497 CCTACTAGATCGCTCCAGACTCGCTTACAGATAGTGTGTACCTCACTGAATATTAT 1556
Qy 241 CCAGTAGAAGATGATGTTATGGAAGGAGAGACTCAATTGAGTCAGTGGCTGATCTGT 300
Db 1557 CCAGTAGAAGATGATGTTATGGAAGGAGAGACTCAATTGAGTCAGTGGCTGATCTGT 1616
Qy 301 GAGCCCACTAAATATCTATGAAGAAACCAACAGACATCTGGATAAATATCCATGACATCT 360
Db 1617 GAGCCCACTAAATATCTATGAAGAAACCAACAGACATCTGGATAAATATCCATGACATCT 1676
Qy 361 TCATGTTTTTCCCAAGTCAAGAGAGAAATGAGTTATTTTTTCCCTCTGAATGCAA 420
Db 1677 TCATGTTTTTCCCAAGTCAAGAGAGAAATGAGTTATTTTTTCCCTCTGAATGCAA 1736
Qy 421 AACAGTTTCCGTCATTATACAAATTTACATCTATTTTAAAGGAAGCAATATAACG 480
Db 1737 AACAGTTTCCGTCATTATACAAATTTACATCTATTTTAAAGGAAGCAATATAACG 1796
Qy 481 ATCCAGTGGTGGCTCGCTGCCAAGTGATTTCAAGTGCTCTATCAAGAGAGATAGC 540
Db 1797 ATCCAGTGGTGGCTCGCTGCCAAGTGATTTCAAGTGCTCTATCAAGAGAGATAGC 1856
Qy 541 AATTACAGTGGTGAATGGGAAGTTCTTGGCGGCGATGGATCTAATATCCAAGTTGATGA 600
Db 1857 AATTACAGTGGTGAATGGGAAGTTCTTGGCGGCGATGGATCTAATATCCAAGTTGATGA 1916
Qy 601 AGTCAGAAGCTGGTATATTTTGAAGGCAACCAAGACTCCCTTTAGAGCATCACCTGTA 660
Db 1917 AGTCAGAAGCTGGTATATTTTGAAGGCAACCAAGACTCCCTTTAGAGCATCACCTGTA 1976
Qy 661 CGTAGTCAGTTACGTAAATCTCTGGAGAGGTGACAAGGCTGACTACCGTGGCTACTCACA 720
Db 1977 CGTAGTCAGTTACGTAAATCTCTGGAGAGGTGACAAGGCTGACTACCGTGGCTACTCACA 2036
Qy 721 TTCTTGCTGCATCAGTCAGCACTGTGACTTCTTTTAAAGTAAGTATAGTAACCAAGAGAA 780
Db 2037 TTCTTGCTGCATCAGTCAGCACTGTGACTTCTTTTAAAGTAAGTATAGTAACCAAGAGAA 2096
Qy 781 TCCACACTGTGTCCCTTTACAAGCTATCAAGTCTCAAGATGACCCAACTTTGCAAAAC 840
Db 2097 TCCACACTGTGTCCCTTTACAAGCTATCAAGTCTCAAGATGACCCAACTTTGCAAAAC 2156
Qy 841 AAAGGAATTTTGGGCCACCATTTTGGATTCAGCAGGTCTCTTCTGACTATACTCTCC 900
Db 2157 AAAGGAATTTTGGGCCACCATTTTGGATTCAGCAGGTCTCTTCTGACTATACTCTCC 2216
Qy 901 AGAAATTTTCTCTTTTGAAGTACTACTGGATTTACATTTGATGGGATGCTCTACAAGCC 960
Db 2217 AGAAATTTTCTCTTTTGAAGTACTACTGGATTTACATTTGATGGGATGCTCTACAAGCC 2276
Qy 961 TCATGATCTACAGCTGGAAAGAAATATCCTACTGTGCTGTTCATATATGCGTGGTCTCA 1020
Db 2277 TCATGATCTACAGCTGGAAAGAAATATCCTACTGTGCTGTTCATATATGCGTGGTCTCA 2336
Qy 1021 ----- 1020
Db 2337 GGTCAGTTGGTGAATAATCGGTTTAAAGAGTCAAGTATTTCCGTTGAATACCCTAGC 2396
Qy 1021 ----- 1020
Db 2397 CTCCTAGTTATGTGTTGTAGTAGATAGACAACAGGGGATCTGTGTCACCGAGGGCTTAA 2456
Qy 1021 -----GGTTGCTATTGCTGGGGCCCCAGTCACTCTGTG 1053
Db 2457 ATTTGAAGCGCCTTTAAATATAAAATGTTGCTATTGCTGGGGCCCCAGTCACTCTGTG 2516
Qy 1054 GATCTTCTATGATACAGGATACACGGAACGTTATATGGGTCAACCCCTGACAGAAATGAACA 1113
```

```
Db 2517 GATCTTCTATGATACAGGATACACGGAACTTTATATGGGTACCCCTGACAGAAATGAACA 2576
Qy 1114 GGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAAAGTTCCCTCTGAACCAAAATCG 1173
Db 2577 GGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAAAGTTCCCTCTGAACCAAAATCG 2636
Qy 1174 TTTACTGCTCTTACATGCGTTTCTCGATGAGAATGCCATTTTGGCACATACCAAGTATAT 1233
Db 2637 TTTACTGCTCTTACATGCGTTTCTCGATGAGAATGCCATTTTGGCACATACCAAGTATAT 2696
Qy 1234 ACTGAGTTTTTTAGTGAGGGCTGGAAAGCCATATGATTTACAGATCTATCTCAGGAGAG 1293
Db 2697 ACTGAGTTTTTTAGTGAGGGCTGGAAAGCCATATGATTTACAGATCTATCTCAGGAGAG 2756
Qy 1294 ACACAGCATAAAGAGTTTCCCTGAATTCGGGAGAAACATTATGAACCTGTCATCTTTGCACTACCT 1353
Db 2757 ACACAGCATAAAGAGTTTCCCTGAATTCGGGAGAAACATTATGAACCTGTCATCTTTGCACTACCT 2816
Qy 1354 TCAGRAAAACCTTTGGATCAGTATTCGCTCTCTAAAGTGATATAATTTTGACCTGTGTA 1413
Db 2817 TCAGRAAAACCTTTGGATCAGTATTCGCTCTCTAAAGTGATATAATTTTGACCTGTGTA 2876
Qy 1414 GAACTCTCTGGTATACACTGGCTATTTACCAAAATGAGAGGTTTAAATCAACAGAAACA 1473
Db 2877 GAACTCTCTGGTATACACTGGCTATTTACCAAAATGAGAGGTTTAAATCAACAGAAACA 2936
Qy 1474 CAGAAATTGATCATCAATTTTGATACCTGCCATGTAACATCTCTCTGAAAAATAAATGT 1533
Db 2937 CAGAAATTGATCATCAATTTTGATACCTGCCATGTAACATCTCTCTGAAAAATAAATGT 2996
Qy 1534 GGTGCCATGACAGGGGCTACGGTTTGGTAGTAATCTAAATACCTTAAACCCCATGCTC 1593
Db 2997 GGTGCCATGACAGGGGCTACGGTTTGGTAGTAATCTAAATACCTTAAACCCCATGCTC 3056
Qy 1594 AAAATCAATGATACATATTTCTTGAGAGAGCCAGCAATACCAATAGAAATTAATAAAAA 1653
Db 3057 AAAATCAATGATACATATTTCTTGAGAGAGCCAGCAATACCAATAGAAATTAATAAAAA 3116
Qy 1654 AAAAAAAAA 1662
Db 3117 AAAAAAAAA 3125

RESULT 9
AR631287 LOCUS AR631287 4676 bp DNA linear PAT 14-FEB-2005
DEFINITION Sequence 20 from patent US 6844180.
ACCESSION AR631287
VERSION AR631287.1 GI:59770934
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 4676)
AUTHORS Oi,S., Akineanya,K.O., Riviere,P.J.M. and Junien,J.-L.
TITLE Serine protease genes related to DPPIV
JOURNAL Patent: US 6844180-A 20 18-JAN-2005;
Ferring BV;;
NLX;

FEATURES
source Location/Qualifiers
1..4676
/mol_type="genomic DNA"

ORIGIN

Query Match 89.9%; Score 1501; DB 6; Length 4676;
Best Local Similarity 91.5%; Pred.No.0;
Matches 1663; Conservative 0; Mismatches 5; Indels 149; Gaps 1;

Qy 1 AACAGGTACAGCAAAATCCCTAAAGTCACCTTTTAAGATGTGAGAAATTAATGATGCTGA 60
Db 1164 AACAGGTACAGCAAAATCCCTAAAGTCACCTTTTAAGATGTGAGAAATTAATGATGCTGA 1223
```

Qy	61	AGAGGATCATAGATGTCATAGATAAAGAACTAAATCAACCTTTTGGAGATCTTATTTGA	120
Db	1224	AGGAAGGATCATAGATGTCATAGATAAAGAACTAAATCAACCTTTTGGAGATCTTATTTGA	1283
Qy	121	AGGAGTTGAATATATTTGCCAGAGCTGGATGGACTCTCTGAGGGGAAATATGCTTGGTCCAT	180
Db	1284	AGGAGTTGAATATATTTGCCAGAGCTGGATGGACTCTCTGAGGGGAAATATGCTTGGTCCAT	1343
Qy	181	CTTACTAGATCGCTCCAGAGCTCGCCTACAGATAGTGTGTGATCTCACCTGAAATATTATTAT	240
Db	1344	CTTACTAGATCGCTCCAGAGCTCGCCTACAGATAGTGTGTGATCTCACCTGAAATATTATTAT	1403
Qy	241	CCAGTAGAAGATGATGTTATGGAAGGCGAGAGCTCATTTGAGTCAGTGCCTCGATTCCTGT	300
Db	1404	CCAGTAGAAGATGATGTTATGGAAGGCGAGAGCTCATTTGAGTCAGTGCCTCGATTCCTGT	1463
Qy	301	GAGCCACATAATTACTATGAAGAAACAAACAGACATCTGGATAAATATCCATGACATCTT	360
Db	1464	GAGCCACATAATTACTATGAAGAAACAAACAGACATCTGGATAAATATCCATGACATCTT	1523
Qy	361	TCATGTTTTTCCCAAAGTTCACGAAGGAGGAAATTTGAGTTTTATTTTTTGGCTCTCGAATGCAA	420
Db	1524	TCATGTTTTTCCCAAAGTTCACGAAGGAGGAAATTTGAGTTTTATTTTTTGGCTCTCGAATGCAA	1583
Qy	421	AACAGGTTTCCGTCTATTTATACAAATTTACATCTATTTTAAAGGAAAGCAAATATAAACG	480
Db	1584	AACAGGTTTCCGTCTATTTATACAAATTTACATCTATTTTAAAGGAAAGCAAATATAAACG	1643
Qy	481	ATTCAGTGTGGGCTGCCTGCTCCAAGTGATTTCAAGTGTCTCTATCAAGAGGAGATAGC	540
Db	1644	ATTCAGTGTGGGCTGCCTGCTCCAAGTGATTTCAAGTGTCTCTATCAAGAGGAGATAGC	1703
Qy	541	AATTACCAAGTGGTGAATGGGAAGTTCTTTGGCCGGCATGGATCTAATATCTCAAAGTTGATGA	600
Db	1704	AATTACCAAGTGGTGAATGGGAAGTTCTTTGGCCGGCATGGATCTAATATCTCAAAGTTGATGA	1763
Qy	601	AGTCAGAAGGCTGGTATATTTTGAAGGCAACAAAGACTCCCTTTTAGAGCATCACTGTGA	660
Db	1764	AGTCAGAAGGCTGGTATATTTTGAAGGCAACAAAGACTCCCTTTTAGAGCATCACTGTGA	1823
Qy	661	CGTAGTCAGTTACGTAATATCTCGAGAGGTTGACAAGGCTGACTGACCGTGGCTACTCACA	720
Db	1824	CGTAGTCAGTTACGTAATATCTCGAGAGGTTGACAAGGCTGACTGACCGTGGCTACTCACA	1883
Qy	721	TTCTTGCTGCATCAGTCAGCACTGTGACTCTTTTATAAGTAAGTATAGTAACACAGAAGAA	780
Db	1884	TTCTTGCTGCATCAGTCAGCACTGTGACTCTTTTATAAGTAAGTATAGTAACACAGAAGAA	1943
Qy	781	TCCACATGTGTGTCCCTTTTACAAGCTATCAAGTCTCTGAAGATGCCAACTTGCAGAAAC	840
Db	1944	TCCACATGTGTGTCCCTTTTACAAGCTATCAAGTCTCTGAAGATGCCAACTTGCAGAAAC	2003
Qy	841	AAAGGAATTTGGGCCACCATTTTGGATTACAGCAGTCTCTCTCTGACTATACTCTCTCC	900
Db	2004	AAAGGAATTTGGGCCACCATTTTGGATTACAGCAGTCTCTCTCTGACTATACTCTCTCC	2063
Qy	901	AGAAATTTTCTCTTTTGAAGTACTACTGCAATTTACATTTGATGGATGCTCTACAAGCC	960
Db	2064	AGAAATTTTCTCTTTTGAAGTACTACTGCAATTTACATTTGATGGATGCTCTACAAGCC	2123
Qy	961	TCATGATCTACAGCTGGAAAGAAATATCTCTGCTGTGTTTCATATATGGTGGTC----	1016
Db	2124	TCATGATCTACAGCTGGAAAGAAATATCTCTGCTGTGTTTCATATATGGTGGTGGGT	2183
Qy	1017	-----CTCAGGTTGCTATTTGCTGGGGCCCCAGTCACTCTG	1051
Db	2184	CAAAATAGAAATTGACGATCAGGTGGAGGACTCCAAATATCTAGCTTCTCGATATGATTTCT	2243
Qy	1017	-----1016	
Db	2244	ATTGACTTAGATCTGTGTGGGCATCCAGGCTGGTCTCTATGAGGATACCTCTCCCTGATG	2303
Qy	1017	-----CTCAGGTTGCTATTTGCTGGGGCCCCAGTCACTCTG	1051

COMMENT	OS Homo sapiens (human)	QY	721	TTCTTGCTGCATCAGTCAGCAGCTGTGACTTCTTTATAGTAAGTATAGTAAACCAAGAA	780
	PN JP 2002191363-A/11844				
	PD 09-JUL-2002	Db	1089	TTCTTGCTGCATCAGTCAGCAGCTGTGACTTCTTTATAGTAAGTATAGTAAACCAAGAA	1148
	PF 28-JUL-2000 JP 2000280990				
	PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU	QY	781	TCACACTGTGTGTCTCTTTTCAAGCTATCAAGTCTCTGGAAGATGACCAACTTGCACAAAC	840
	PI SAITO,				
	PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,	Db	1149	TCACACTGTGTGTCTCTTTTCAAGCTATCAAGTCTCTGGAAGATGACCAACTTGCACAAAC	1208
	PI KEIICHI NAGAI, TETSUJI OTSUKI				
	PC C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC	QY	841	AAAGGAATTTTGGGCCACCATTTCGATTTCAGCAGGTCTCTTCTGCTGACTATATCTCTCC	900
	10,				
	PC C12P21/02, C12Q1/68, C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC	Db	1209	AAAGGAATTTTGGGCCACCATTTCGATTTCAGCAGGTCTCTTCTGCTGACTATATCTCTCC	1268
	Primer for synthesizing full-length cDNA and use thereof FH Key				
	Location/Qualifiers	QY	901	AGAAATTTTCTCTCTTTTGAAGTACTACTGGATTTCATTTGATGGGATGCTCTCAAGCC	960
	(16) (1911).				
FEATURES	CDS	Db	1269	AGAAATTTTCTCTCTTTTGAAGTACTACTGGATTTCATTTGATGGGATGCTCTCAAGCC	1328
source	Location/Qualifiers	QY	961	TCATGATCTACAGCTGGAAAGAAATNTCTACTGTGCTGTTTCATATATGTTGCTCTCA	1020
	1..2161				
	/organism="Homo sapiens"	Db	1329	TCATGATCTACAGCTGGAAAGAAATNTCTACTGTGCTGTTTCATATATGTTGCTCTCA	1388
	/mol_type="genomic DNA"				
	/db_xref="taxon:9606"	QY	1021	-----	1020
ORIGIN		Db	1389	GGTGCAAGTTGGTGAATAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTTAGC	1448
	Query Match 89.1%; Score 1487.4; DB 6; Length 2161;	QY	1021	-----	1020
	Best Local Similarity 91.7%; Pred. No. 0;	Db	1449	CTCTCTAGGTTATGTGGTTGTAGTAGATAGACAAACAGGGGATCCTGTCTACCGAGGCTTAA	1508
	Matches 1645; Conservative 0; Mismatches 1; Indels 147; Gaps 1;	QY	1021	-----	1020
		Db	1509	ATTTGAAGCGCGCTTTAAATATATAAAATGTTGCTTATTCGCGGCCCGAGTCACTCTGTG	1568
		QY	1054	GATCTTCTATGATACAGGATACACGGAAGTTTATATGTTGCTGCTGACCAAGAAATGAACA	1113
		Db	1569	GATCTTCTATGATACAGGATACACGGAAGTTTATATGTTGCTGCTGACCAAGAAATGAACA	1628
		QY	1114	GGGCTATTTACTTAGGATCTGTGGCCATGCAACGAGAAAAGTTTCCCTCTGAACCAAAATCG	1173
		Db	1629	GGGCTATTTACTTAGGATCTGTGGCCATGCAACGAGAAAAGTTTCCCTCTGAACCAAAATCG	1688
		QY	1174	TTTACTGCTTTACATGCTGTTTCTGGATGGAATGCTCAATTTTGACATACACAGTATAT	1233
		Db	1689	TTTACTGCTTTACATGCTGTTTCTGGATGGAATGCTCAATTTTGACATACACAGTATAT	1748
		QY	1234	ACTGAGTTTATTTAGTCAGGCTGGAAAGCCATATGATTTTACAGATCTATCTCTCAGGAGAG	1293
		Db	1749	ACTGAGTTTATTTAGTCAGGCTGGAAAGCCATATGATTTTACAGATCTATCTCTCAGGAGAG	1808
		QY	1294	ACACAGCATAAAGAGTTTCTTGAATCGGGAACAATTTATGAACCTGCAATCTTTTGCACCT	1353
		Db	1809	ACACAGCATAAAGAGTTTCTTGAATCGGGAACAATTTATGAACCTGCAATCTTTTGCACCT	1868
		QY	1354	TCAGAAAACCTTTGGATACGTTATGCTCTCTTAAAGATGATATAATTTTGACCTGTGTA	1413
		Db	1869	TCAGAAAACCTTTGGATACGTTATGCTCTCTTAAAGATGATATAATTTTGACCTGTGTA	1928
		QY	1414	GAATCTCTGGTATACATCTGGCTTATTAACCAAAATGAGGAGTTTAAATCAACAGAAAACA	1473
		Db	1929	GAATCTCTGGTATACATCTGGCTTATTAACCAAAATGAGGAGTTTAAATCAACAGAAAACA	1988
		QY	1474	CAGAAATTTGATCATCATCTTTTGTATCTCCCATGTAACATCTTCTCTGAAAATAAATGT	1533
		Db	1989	CAGAAATTTGATCATCATCTTTTGTATCTCCCATGTAACATCTTCTCTGAAAATAAATGT	2048
		QY	1534	GGTGCCATGCGGGGCTACGGTTTGTAGTAATCTTAATACCTTAAACCCCAATGCTC	1593
		Db	2049	GGTGCCATGCGGGGCTACGGTTTGTAGTAATCTTAATACCTTAAACCCCAATGCTC	2108
		QY	1594	AAAATCAATGATACATATTTCTTGAGAGACCCAGCAATACCATAGAATTAAT	1646
		Db	2109	AAAATCAATGATACATATTTCTTGAGAGACCCAGCAATACCATAGAATTAAT	2161

QY 1354 TCAAGAAAACCTTGGATCAGGTAATGCTGCTCTAAAAGTGATATAAATTTTGACCTGTGTA 1413
Db 1869 TCAAGAAAACCTTGGATCAGGTAATGCTGCTCTAAAAGTGATATAAATTTTGACCTGTGTA 1928
QY 1414 GAATCTCTGGTATACACATGCGGTATTAACCAATGAGGAGTTTAATCAACGAAACA 1473
Db 1929 GAATCTCTGGTATACACATGCGGTATTAACCAATGAGGAGTTTAATCAACGAAACA 1988
QY 1474 CAGAAATGATCAATTTTGATACCTGCGCATGTAACATCTACTCTCTGAAATAAATGT 1533
Db 1989 CAGAAATGATCAATTTTGATACCTGCGCATGTAACATCTACTCTCTGAAATAAATGT 2048
QY 1534 GTGGCCATCAGGGGTCTACGGTTTGTGTAGTAAATCTTAACCCCATGCTC 1593
Db 2049 GTGGCCATCAGGGGTCTACGGTTTGTGTAGTAAATCTTAACCCCATGCTC 2108
QY 1594 AATCAATGATACATATTTCTGAGACCCAGCAATACCATGAAGATTACT 1646
Db 2109 AATCAATGATACATATTTCTGAGACCCAGCAATACCATGAAGATTACT 2161

RESULT 13
AK027826 2161 bp mRNA linear PRI 30-JAN-2004
LOCUS Homo sapiens cDNA FLJ14920 fis, clone PLACE1007416, weakly similar
DEFINITION to DIPEPTIDYL PEPTIDASE IV (EC 3.4.14.5).
ACCESSION AK027826
VERSION AK027826.1 GI:14042789
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.

REFERENCE 1
Ota,T., Suzuki,Y., Nishikawa,T., Otsuki,T., Sugiyama,T., Irie,R.,
Wakabatu,A., Hayashi,K., Sato,H., Nagai,K., Kimura,K., Makita,H.,
Sekine,M., Oabayashi,M., Nishi,T., Shibahara,T., Tanaka,H.,
Ishii,S., Yamamoto,J., Saito,K., Kawai,Y., Isono,Y., Nakamura,Y.,
Nagahari,K., Murakami,K., Yasuda,T., Iwayanagi,T., Wagatsuma,M.,
Shiratori,A., Sudo,H., Hosoiri,T., Kaku,Y., Kodaira,H., Kondo,H.,
Sugawara,M., Takahashi,M., Kanda,K., Kamehara,K., Katsuta,N., Sato,K.,
Kikkawa,E., Omura,Y., Abe,K., Kamihara,K., Katsuta,N., Sato,K.,
Tanikawa,M., Yamazaki,M., Ninomiya,K., Ishibashi,T., Yamashita,H.,
Murakawa,K., Fujimori,K., Tanai,H., Kimata,M., Watanabe,M.,
Hiraoka,S., Chiba,Y., Ishida,S., Ono,Y., Takiguchi,S., Watanabe,S.,
Yosida,M., Houta,T., Kusano,J., Kanehori,K., Takahashi-Fujii,A.,
Hara,H., Tanase,T., Nomura,Y., Togiya,S., Komai,F., Hara,R.,
Takeuchi,K., Arita,M., Imose,N., Musashino,K., Yuuki,H., Oshima,A.,
Sasaki,N., Aotaka,S., Yoshikawa,Y., Matsunawa,H., Ichihara,T.,
Shiohata,N., Sano,S., Moriya,S., Momiyama,H., Satoh,N., Takami,S.,
Terahima,Y., Suzuki,O., Nakagawa,S., Senoh,A., Mizoguchi,H.,
Goto,Y., Shimizu,F., Wakebe,H., Hishigaki,H., Watanabe,T.,
Sugiyama,A., Takemoto,M., Kawakami,Y., Yamazaki,M., Watanabe,K.,
Kumagai,A., Itakura,S., Fukuzumi,Y., Fujimori,Y., Komiyama,M.,
Tashiro,H., Tanigami,A., Fujiwara,T., Ono,T., Yamada,K., Fujii,Y.,
Ozaki,K., Hirao,M., Ohmori,Y., Kawabata,A., Hikiji,T., Kobatake,N.,
Inagaki,H., Ikema,Y., Okamoto,S., Okitani,R., Kawakami,T.,
Noguchi,S., Itoh,T., Shigetake,K., Senba,T., Matsumura,K.,
Nakajima,Y., Mizuno,T., Morinaga,M., Sasaki,M., Togashi,T.,
Oyama,M., Hata,H., Watanabe,M., Komatsu,T., Mizushima-Sugano,J.,
Satoh,T., Shirai,Y., Takahashi,Y., Nakagawa,K., Okumura,K.,
Nagata,T., Nomura,N., Kikuchi,H., Masuho,Y., Yamashita,R.,
Nakai,K., Yada,T., Nakamura,Y., Ohara,O., Isogai,T. and Sugano,S.
Complete sequencing and characterization of 21,243 full-length
human cDNAs
Nat. Genet. 36 (1), 40-45 (2004)

JOURNAL
PUBMED
REFERENCE 2
AUTHORS
Isogai,T., Ota,T., Hayaishi,K., Sugiyama,T., Otsuki,T., Suzuki,Y.,
Nishikawa,T., Nagai,K., Sugano,S., Shiratori,A., Sudo,H.,
Wagatsuma,M., Hosoiri,T., Kaku,Y., Kodaira,H., Kondo,H.,

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT

FEATURES
source

CDS

ORIGIN

Query Match 89.1%; Score 1487.4; DB 8; Length 2161;
Best Local Similarity 91.7%; Pred. No. 0;
Matches 1645; Conservative 0; Mismatches 1; Indels 147; Gaps 1;
QY 1 AACAGGTACAGCAATCCTAAAGTCACTTTTAAGATGTCAGAAATAATGATGATGCTGA 60
Db 369 AACAGGTACAGCAATCCTAAAGTCACTTTTAAGATGTCAGAAATAATGATGATGCTGA 428
QY 61 AGGAAGGATCATAGATGTCATAGATAGGAACCTTAATCAACCTTTGAGATTCATTATGGA 120
Db 429 AGGAAGGATCATAGATGTCATAGATAGGAACCTTAATCAACCTTTGAGATTCATTATGGA 488
QY 121 AGGAGTTGAATATATTCGCCAGACTGGATCGACTCTGAGGAGAAATATGCTTGTCAT 180
Db 489 AGGAGTTGAATATATTCGCCAGACTGGATCGACTCTGAGGAGAAATATGCTTGTCAT 548
QY 181 CCTACTAGATCGCTCCAGACTCGCTCTAGATAGTGTGTGATCTCACCTGGAATATTTAT 240
Db 549 CCTACTAGATCGCTCCAGACTCGCTCTAGATAGTGTGTGATCTCACCTGGAATATTTAT 608
QY 241 CCCAGTAGAAGATGATGTTATGGAAGGCCAGAGACTCATTGAGTCAGTCGCTGATCTGT 300
Db 609 CCCAGTAGAAGATGATGTTATGGAAGGCCAGAGACTCATTGAGTCAGTCGCTGATCTGT 668
QY 301 GACCCCAATATATCTATCAAGAAACACAGACATCTCGATAAATATATCATGACATCTT 360
|||||

Sugawara,M., Takahashi,M., Chiba,Y., Ishida,S., Murakawa,K.,
Ono,Y., Takiguchi,S., Watanabe,S., Kimura,K., Murakami,K.,
Ishii,S., Kawai,Y., Saito,K., Yamamoto,J., Wakamatsu,A.,
Nakamura,Y., Nagahari,K., Masuho,Y., Ninomiya,K. and Iwayanagi,T.
NEDO human cDNA sequencing project
Unpublished
3 (bases 1 to 2161)
Isogai,T. and Otsuki,T.
Direct Submission
Submitted (10-MAY-2001) Takao Isogai, Helix Research Institute,
Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing;
Research Association for Biotechnology; cDNA library construction;
5'- & 3'-end one pass sequencing and clone selection; Helix
Research Institute (supported by Japan Key Technology Center etc.)
and Department of Virology, Institute of Medical Science,
University of Tokyo.
Location/Qualifiers
1. .2161
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="PLACE1007416"
/tissue_type="placenta"
/clone_lib="PLACE1"
/notes="cloning vector: pME18SFL3"
16. .1914
/note="unnamed protein product"
/codon_start=1
/protein_id="BAB55395.1"
/db_xref="GI:14042790"
/translation="MDPKLCPADPDWIAFIHSNDIWNISNIVTREERTLYVHNELANN
EEDARGAVATFVLOEFDYRGVWVKPKAETPSGGKIRILRYEENDESEVEIHHVT
SPMLTRRADSFYKPTGTANKPTFKMSEIMDAEGRIIDVIDKELIQPEILPEGV
EYARAGWTEPGKYAWSILLDSQRLQIVLISPELFIPVEDDWMEROLLESVPDSV
TPILIIYEETDWINIHDFVPPQSHEEIEIFIPASECKTGRFRLYKILKESKY
KRGSGLPAPSPFKPIKEEIAITGEWELRHGNSIQVDEVRLLYKIFEGTKDSPLE
HHLVYVYVNVGVEVRLTRDYSHSQCCISQHCDFPISKYSQKPHCVSLYKLJASPED
DPYCKTEFWATILDSAGPLPYTPPEIFSPSTTGTLYGLMKPHCLQPKGYKPTV
LFYGGPQVLVNNREKGVKVFRLATLASLGVVVVINDRSGHRLKFECAFKYKVM
AIGAPVTLWIFVDTGYTRYMGHPDQEQYGLSVAMQAEKFPSEPNRLLLHGLF
DENVFAHTSILLSLFVLRAGKPYDQIYPPQERHSIRVPESGEHVELHLHLYQENLGS
RIAAKVI"

TFLEQAGSGLYHVKDGGPOGFTQOPLRPNLVETSCPNIRMBPKLCPADPDWIAPIHNS
DIWISNIVTREERLTYVHNELANEEDARSAGVATFVLQEEFDRIQSYWCPRAEIT
PSGKILIRLILENDESEVEI IHVTSPLMETRRADSPRYKPTGTANPKVTFPMSEIMI
DAEGRIIDVLDKELIOPFEILFEGVEYIARAGWTPGKYAWSILLDRSQVRLQIWLIS
PELFI PVEDDWMERQRLIESVDSVPLI IYEETTDIWINIHDIPHVPQSHHEIEIF
IFASECKTGFRLYKITSILKESKYKSSGSLPAPSPDFKCPKBEI AITSGEWELGR
HGSNIQVDEVRILVYPEGPKDSPLEHLHVUSYVNPGEVTRBLTDRGYSHSCCISQHD
FFTSKYSNQKPHCVSLYKLSPPEDDPTCKTKFEPWATILDSAGPLPDYTPPEIPSPES
TVVITDNGSLYKHLDPQPKGYPTVLFYIGGPQVQLVNNRFGKVKYFRNLANTLASLGV
WVINDRSGHRLKFEFAGFYKMGQIEIDQDQVEGLQYLASRYDFIDLDRVGIHGWSY
GCVLSMALMQRSDI FRVAIAGAPVTWTFYDGYTERYMGHPDQNEOGYVILGSVAMQ
AEKFPSPNRLILLHGFLENVHFAHTSILLSFLVRAGKPYDLQIYQBERHSIRVPS
GEHYELHLLHYLQENLGSRIAALKV"

ORIGIN

Query Match 81.1%; Score 1354; DB 8; Length 3127;

Best Local Similarity 84.7%; Pred. No. 9.2e-296;

Matches 1664; Conservative 0; Mismatches 0; Indels 300; Gaps 1;

QY	1	AACAGGTACAGCAATCCTAAAGTCACCTTTAAGATGTCAGAAATAATGATGATGCTGA	60
Db	1164	AACAGGTACAGCAATCCTAAAGTCACCTTTAAGATGTCAGAAATAATGATGATGCTGA	1223
QY	61	AGGAAGCATCATAGATGTCATAGATAGGAATCAATTCACCTTTTGAGATTCCTATTGA	120
Db	1224	AGGAAGCATCATAGATGTCATAGATAGGAATCAATTCACCTTTTGAGATTCCTATTGA	1283
QY	121	AGGAGTTGAATATATTGCGAGACTGGATGGACTCTCTGAGGGAAATAATGCTGGTCCAT	180
Db	1284	AGGAGTTGAATATATTGCGAGACTGGATGGACTCTCTGAGGGAAATAATGCTGGTCCAT	1343
QY	181	CCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGTGATCTCACCTGAATATTAT	240
Db	1344	CCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGTGATCTCACCTGAATATTAT	1403
QY	241	CCCACTAGATAGATGTTATGGAAGGAGGAGACTCATTGAGTCAGTGCCTGATTCGT	300
Db	1404	CCCACTAGATAGATGTTATGGAAGGAGGAGACTCATTGAGTCAGTGCCTGATTCGT	1463
QY	301	GAGGCCACTAATTTATCTATGAAGAAACAACAGACATCTGGATAATATCCATGACATCTT	360
Db	1464	GAGGCCACTAATTTATCTATGAAGAAACAACAGACATCTGGATAATATCCATGACATCTT	1523
QY	361	TCATGTTTTTCCCAAGTCACGAAGAGAAATTGAGTTATTTTGGCTCTGAATGCAA	420
Db	1524	TCATGTTTTTCCCAAGTCACGAAGAGAAATTGAGTTATTTTGGCTCTGAATGCAA	1583
QY	421	AACAGGTTTCCGTCATTTATACAAATTTACATCTATTTTAAAGGAAGCAATATAACG	480
Db	1584	AACAGGTTTCCGTCATTTATACAAATTTACATCTATTTTAAAGGAAGCAATATAACG	1643
QY	481	ATCCAGTGGTGGCTGCTGCTCCAAGTATTTCAAGTGTCTATCAAGAGGAGATAGC	540
Db	1644	ATCCAGTGGTGGCTGCTGCTCCAAGTATTTCAAGTGTCTATCAAGAGGAGATAGC	1703
QY	541	AAATTACAGTGGTGAATGGGAAGTCTTGGCCGGCATGGATCTAATATCCAAAGTTGATGA	600
Db	1704	AAATTACAGTGGTGAATGGGAAGTCTTGGCCGGCATGGATCTAATATCCAAAGTTGATGA	1763
QY	601	AGTCAGAGGCTGGTATATTTTGAAGGCAACAAAGACTCCCTTTTGAAGCATACCTGTGA	660
Db	1764	AGTCAGAGGCTGGTATATTTTGAAGGCAACAAAGACTCCCTTTTGAAGCATACCTGTGA	1823
QY	661	CGTAGTCAGTTAGTAAATCCCTGGAGAGGTGACAGGCTGACTGACCGTGGCTACTACA	720
Db	1824	CGTAGTCAGTTAGTAAATCCCTGGAGAGGTGACAGGCTGACTGACCGTGGCTACTACA	1883
QY	721	TTCTTGCTGATCAGTCAGTGTGACTTCTTTTATAAGTAAAGTATAGTACCAAGAGAA	780
Db	1884	TTCTTGCTGATCAGTCAGTGTGACTTCTTTTATAAGTAAAGTATAGTACCAAGAGAA	1943
QY	781	TCACACTGTGTGTCCTTTTACAGCTATCAAGTCTCTGAGATGACGCCAATCTTGCANAAC	840

Db	1944	TCACACTGTGTGTCCTTTTACAAGCTATCAAGTCTCTGAAGATGACCCAACTTGCANAAC	2003
QY	841	AAAGGAATTTTGGGCCACCACTTTTGGATTCAGCAGGTCTCTTCTGACTATATCTCTCC	900
Db	2004	AAAGGAATTTTGGGCCACCACTTTTGGATTCAGCAGGTCTCTTCTGACTATATCTCTCC	2063
QY	901	AGAAATTTTCTCTTTTGAAGTACTACTGGATTTACATTTGATGGATGCTCTACAAGCC	960
Db	2064	AGAAATTTTCTCTTTTGAAGTACTACTGGATTTACATTTGATGGATGCTCTACAAGCC	2123
QY	961	TCATCATCTACAGCTGGAAGAAATATCTCTACTGTGCTGCTTCATATATGGTGGTCTCA	1020
Db	2124	TCATCATCTACAGCTGGAAGAAATATCTCTACTGTGCTGCTTCATATATGGTGGTCTCA	2183
QY	1021	-----	1020
Db	2184	GGTGCAGTTGCTGAATAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCCTAGC	2243
QY	1021	-----	1020
Db	2244	CTCTCAGGTTTATGCTGTTGTAGTGATAGACAAACAGGGGATCTCTGTCACCGAGGCTTAA	2303
QY	1021	-----	1020
Db	2304	ATTTGAAGCGCCTTTTAAATATATAAATGGGTCAAAATGACAAATTGACGATCAGGTGGAAGG	2363
QY	1021	-----	1020
Db	2364	ACTCCAAATATCTAGCTTCTCGATATGATTTTCATTGACTTAGATCGTGTGGGCATCCACGG	2423
QY	1021	-----	1020
Db	2424	CTGGTCTTATGGAGGATACCTCTCCCTGATGGCATTAATGACAGAGGTTCAGATATCTTCAG	2483
QY	1021	-----	1020
Db	1021	GGTGTCTATTGCTGGGGCCCACTCTCTGTGGATCTCTTATGATACAGGATACACGCA	1080
Db	2484	GTTGTCTATTGCTGGGGCCCACTCTCTGTGGATCTCTTATGATACAGGATACACGCA	2543
QY	1081	ACGTTTATATGGGTCACTCCCTGACCAAGATGAACAGGCTATTACTTAGGATCTGTGGCCAT	1140
Db	2544	ACGTTTATATGGGTCACTCCCTGACCAAGATGAACAGGCTATTACTTAGGATCTGTGGCCAT	2603
QY	1141	GCAAGCAGAAAGTTTCCCTCTGAAACCAATCGTTTACTGTCTCTTACATGGTTTCTGGA	1200
Db	2604	GCAAGCAGAAAGTTTCCCTCTGAAACCAATCGTTTACTGTCTCTTACATGGTTTCTGGA	2663
QY	1201	TGAGATGTCCATTTTGCACATACCAAGTATTTACTGAGTCTTTTATGACGGCTGGAA	1260
Db	2664	TGAGATGTCCATTTTGCACATACCAAGTATTTACTGAGTCTTTTATGACGGCTGGAA	2723
QY	1261	GCCATATGATTTACAGATCTATCTCTAGGAGAGACACAGCATAAAGATTCCTGAATCGGG	1320
Db	2724	GCCATATGATTTACAGATCTATCTCTAGGAGAGACACAGCATAAAGATTCCTGAATCGGG	2783
QY	1321	AGAACTATTGAACTGCACTTTTGGCACTACCTTCAAGAAAACTTTGGATCACTGATTGC	1380
Db	2784	AGAACTATTGAACTGCACTTTTGGCACTACCTTCAAGAAAACTTTGGATCACTGATTGC	2843
QY	1381	TGCTCTAAAAGTATATAATTTTGAACCTGTGTAGAACTCTCTGGTATATACACTGGCTATT	1440
Db	2844	TGCTCTAAAAGTATATAATTTTGAACCTGTGTAGAACTCTCTGGTATATACACTGGCTATT	2903
QY	1441	AACCAATATGAGGAGTTTAAATCAACAGAAAAACAGAAATTTGATCATCATTTTGCATACC	1500
Db	2904	AACCAATATGAGGAGTTTAAATCAACAGAAAAACAGAAATTTGATCATCATTTTGCATACC	2963
QY	1501	TGCCATGTAAACATCTACTCTCTGAAAAATAATGTGGTCCATGCAGGGGTCTACCGTTTGT	1560
Db	2964	TGCCATGTAAACATCTACTCTCTGAAAAATAATGTGGTCCATGCAGGGGTCTACCGTTTGT	3023
QY	1561	GGTAGTAAATCTAATACCTTAAACCCACATGCTCAAAATCAAAATGATACATATTTCTGAGA	1620
Db	3024	GGTAGTAAATCTAATACCTTAAACCCACATGCTCAAAATCAAAATGATACATATTTCTGAGA	3083

QY	1621	GACCCAGCAATACCAAGAAATTTACTAATAAATAAATAAATAAATAA 1664	Db	1419	CCCAGTAGAAGATGATGTTTATGGAAGGCAGAGACTCATTGAGTCAGTGCCTGATTCTGT 1478
Db	3084	GACCCAGCAATACCAAGAAATTTACTAATAAATAAATAAATAAATAA 3127	QY	301	GACGCCACTAATTTATCTATGAAGAAACACAGACATCTGGATAAATATCCATGACATCTT 360
RESULT 15					
AX354793	AX354793	3143 bp	DNA	linear	PAT 06-FEB-2002
LOCUS	Sequence	1 from Patent WO0179473.			
DEFINITION	AX354793				
ACCESSION	AX354793.1	GI:18619526			
VERSION					
KEYWORDS					
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE					
AUTHORS	Meyers, R.A. and Williamson, M.				
TITLE	21953, a human prollyl oligopeptidase family member and uses thereof				
JOURNAL	Patent: WO 0179473-A 1 25-OCT-2001;				
FEATURES	Millennium Pharmaceuticals, Inc. (US)				
source	Location/Qualifiers				
	1..3143				
	/organism="Homo sapiens"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:9606"				
	229..2877				
	/note="unnamed protein product"				
	/codon_start=1				
	/protein_id="CAD22990.1"				
	/db_xref="GI:18619527"				
	/translation="MAAMETEQLGVEIFETADCEENIESQDRPKLEPFVVERYSWSQ				
	LKLLADATRYHYGMMAKAPDFMEVKNRDPDGHSDRIIYLAGMNGRENTLIFYSEI				
	PKTINRAAIVMLSWPLLDLFOATLDYDMSREBELLRKRRIKRGIVGIASIVYHQSSG				
	TLFQAGSGIVHYKGGPGQFTQQPLRNVLVETSCPNIRMDPKLCPADPQWIAPIHSN				
	DWISINVTRERRITYVHNELANMEEDARSAGVATFVLOEEDRVSQYWCPKAET				
	PSGCKLRLIYXENDESEVEIHHVSPMLETRADSPRYKPTGANKPVTFKMSLMI				
	DAGRLIDVIDKELQPELIFEVEYIARAGMTPEGIKAWMSILLDRSQTRQLIVLIS				
	PELFIVEDVMERQLIESVPSVTPLIYYEETDIWINIHIFVFPQSHHEEIEF				
	IFASECKTFRHLXITISLKESKYRSGGLPAPDFKCPKEEAITSGEWEVLEF				
	HGNSIQVDRNRPCHVSLYKLSPPEDPTCKFEWATILDSAGPLPDYTPPELFSFES				
	FTFKSYGNKPNCHVSLYKLSPPEDPTCKFEWATILDSAGPLPDYTPPELFSFES				
	TTGFTLYGLMYLPHDLQPKKYPTVLFITYGQVQLVNNRFKGVKPYRLNLTSLGTV				
	VVVIDNRGSGHRLKFEAGAFYKMGQIEIDQVEGLQLASRYDFIDLRVGIHGWSY				
	GGYLSLMLMQRSDFRVAIAGAPVTLMFYDTGTYERYMHPDQNEQGYLYLGSVAMQ				
	AEKFPSEPRLLLHGFLDENVHFAHTSILLSFLVRAGKPYDLQIYPOERHSIRVPES				
	GEHYELHLHYLQENLGSRIAALKVI"				
ORIGIN					
Query Match	81.1%;	Score 1353.4;	DB 6;	Length 3143;	
Best Local Similarity	84.7%;	Pred. No. 1.3e-295;			
Matches 1664;	Conservative 0;	Mismatches 1;	Indels 300;	Gaps 1;	
QY	1	ACAGGTACAGCAATCCTAAGATCACTTTTAAGATGTCAGAAATATGATGATGCTGA 60			
Db	1179	ACAGGTACAGCAATCCTAAGATCACTTTTAAGATGTCAGAAATATGATGATGCTGA 1238			
QY	61	AGGAAGATCATAGATGTCATAGATAAGGAACATAATCAACCTTTTGAGATCTTATTGA 120			
Db	1239	AGGAAGATCATAGATGTCATAGATAAGGAACATAATCAACCTTTTGAGATCTTATTGA 1298			
QY	121	AGGAGTTGAATATATTGCCAGAGCTGGATGGACTCCTCAGGGAATAATATGCTGGTCCAT 180			
Db	1299	AGGAGTTGAATATATTGCCAGAGCTGGATGGACTCCTCAGGGAATAATATGCTGGTCCAT 1358			
QY	181	CCTACTAGATCGCTCCAGACCTCGCTACAGATAGTGTGTGCTCACTCACTGAATTTATTAT 240			
Db	1359	CCTACTAGATCGCTCCAGACCTCGCTACAGATAGTGTGTGCTCACTCACTGAATTTATTAT 1418			
QY	241	CCCAAGTAGAAGATGATGTTATGGAAGGCAGAGACTCAATTCAGTCAGTGCCTGATTCGT 300			

Qy	1081	ACGTTATATGGTCCACCTGACAGAAATGAACAGGGCTATTACTTAGGATCTGTGGCCAT	1140
Db	2559	ACGTTATATGGGTACCCCTGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCCAT	2618
Qy	1141	GCAAGCAGAAAAGTTCCCTCTGAAACCAAAATCGTTTACTGCTCTTACATGGTTTCTCGGA	1200
Db	2619	GCAAGCAGAAAAGTTCCCTCTGAAACCAAAATCGTTTACTGCTCTTACATGGTTTCTCGGA	2678
Qy	1201	TGAGAAATGTCATTTTGGACATACACAGTATATTTACTGAGTTTCTTACTGAGGGCTGGAAA	1260
Db	2679	TGAGAAATGTCATTTTGGACATACACAGTATATTTACTGAGTTTCTTACTGAGGGCTGGAAA	2738
Qy	1261	GCATATGATTTTACAGATCTATCTCAGGAGAGACAGCATAAAGAGTTCTTGAATCGGG	1320
Db	2739	GCATATGATTTTACAGATCTATCTCAGGAGAGACAGCATAAAGAGTTCTTGAATCGGG	2798
Qy	1321	AGAACATTTATGAATGCACTCTTTTGGCACTACCTTCAAGAAAAACCTTGGATCAAGTATTCG	1380
Db	2799	AGAACATTTATGAATGCACTCTTTTGGCACTACCTTCAAGAAAAACCTTGGATCAAGTATTCG	2858
Qy	1381	TGCTCTAAAGTGATATTAATTTGACCTGTGTAGAACTCTCTGGTATACACTGGCTATTTT	1440
Db	2859	TGCTCTAAAGTGATATTAATTTGACCTGTGTAGAACTCTCTGGTATACACTGGCTATTTT	2918
Qy	1441	AACCAATGAGGAGGTTTAAATCAACAGAAAAACAGAAATTTGATCATCACATTTTGATACC	1500
Db	2919	AACCAATGAGGAGGTTTAAATCAACAGAAAAACAGAAATTTGATCATCACATTTTGATACC	2978
Qy	1501	TGCCATGTAACTCTACTCTGAAATAAATGTGGTGCCATGCAGGGGTCTACGGTTTGT	1560
Db	2979	TGCCATGTAACTCTACTCTGAAATAAATGTGGTGCCATGCAGGGGTCTACGGTTTGT	3038
Qy	1561	GGTAGTAATCTAATACCTTTAACCCCAATGCTCAAAATCAAAATGATACATATTTCTCGAGA	1620
Db	3039	GGTAGTAATCTAATACCTTTAACCCCAATGCTCAAAATCAAAATGATACATATTTCTCGAGA	3098
Qy	1621	GACCCAGCAATACCATTAAGAAATTTACTAAAAAATTTAAAAAATTTAAAAAATTTAAAAA	1665
Db	3099	GACCCAGCAATACCATTAAGAAATTTACTAAAAAATTTAAAAAATTTAAAAAATTTAAAAA	3143

Search completed: May 4, 2006, 18:51:59
Job time : 8334.8 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 4, 2006, 02:50:44 ; Search time 939.948 Seconds
(without alignments)
11834.025 Million cell updates/sec

Title: US-10-825-632-6
Perfect score: 1669
Sequence: 1 aacaggtagcaatccta.....aaaaaaaaaaaaaaaa 1669

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues
Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : N_Geneseq_21.*
- 1: Geneseqn1980s.*
 - 2: Geneseqn1990s.*
 - 3: Geneseqn2000s.*
 - 4: Geneseqn2001as.*
 - 5: Geneseqn2001bs.*
 - 6: Geneseqn2002as.*
 - 7: Geneseqn2002bs.*
 - 8: Geneseqn2003as.*
 - 9: Geneseqn2003bs.*
 - 10: Geneseqn2003cs.*
 - 11: Geneseqn2003ds.*
 - 12: Geneseqn2004as.*
 - 13: Geneseqn2004bs.*
 - 14: Geneseqn2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1669	100.0	1669	4 AAC85696	AAC85696 Nucleotid
2	1659	99.4	2668	6 ABN59775	ABN59775 Novel hum
3	1659	99.4	2842	6 ABN59774	ABN59774 Novel hum
4	1648	94.8	4523	6 ABK83325	ABK83325 cDNA enco
5	1614	96.7	2702	13 ADT04072	ADT04072 Human pro
6	1501	89.9	4676	6 ABK83331	ABK83331 cDNA enco
7	1493	89.5	2952	10 ACA92421	ACA92421 DNA enco
8	1487	89.1	2161	4 AAH15009	AAH15009 Human pro
9	1399	83.8	2349	14 ADV43982	ADV43982 Human psy
10	1353	81.1	3143	6 AAH99934	AAH99934 cDNA enco
11	1350	80.9	2083	7 ADR41222	ADR41222 Human CD-
12	1349	80.9	4829	6 ABK83327	ABK83327 cDNA enco
13	1347	80.7	3120	4 AAC85694	AAC85694 Nucleotid
14	1347	80.7	3120	6 AAD38956	AAD38956 Human dip
15	1344	80.5	3106	6 ABK12892	ABK12892 Human pro
16	1343	80.5	4685	6 ABK83332	ABK83332 cDNA enco
17	1341	80.3	2929	10 ACA92425	ACA92425 DNA enco
18	1337	80.1	2797	12 ADL13374	ADL13374 Human ste
19	1266	75.9	2696	12 ADI16386	ADI16386 Human pro

20	1250.6	74.9	2510	6 AAD23843	Aad23843 Human pro
21	1209.4	72.5	1821	6 ABV76411	Abv76411 Diptetidy
22	1183.6	70.9	2463	10 ACA92424	ACA92424 DNA enco
23	1089	65.2	2649	14 ADV43981	Adv43981 Human psy
24	1085.8	65.1	2649	8 ABX12255	Abx12255 cDNA enco
25	1085.8	65.1	2671	6 ABK83322	Abk83322 cDNA enco
26	1084.4	65.0	2643	6 AAH99935	Aah99935 Coding se
27	1040.6	62.3	2251	10 ADE79035	Ade79035 Human pro
28	881	52.8	1197	4 AAC85695	Aac85695 Nucleotid
29	765.2	45.8	4309	6 ABK83328	Abk83328 cDNA enco
30	636.4	38.1	873	4 AAF81719	Aaf81719 Human pro
31	636	38.1	925	6 ABL90148	AbL90148 Human pol
32	523.6	31.4	587	4 AAH12830	Aah12830 Human CDN
33	493.4	29.6	823	6 ABK30401	Abk30401 Human G-p
34	434.2	26.0	1083	4 AAC85697	Aac85697 Nucleotid
35	321	19.2	2751	6 AAD38311	Ada38311 Murine di
36	321	19.2	3287	6 AAD38955	Aad38955 Alternati
37	313	18.8	2495	6 AAD38957	Aad38957 Human dip
38	313	18.8	2617	6 ABK83323	Abk83323 cDNA enco
39	313	18.8	2660	6 ADI16689	Adi16689 Human NOV
40	313	18.8	2660	6 ADI16687	Adi16687 Human NOV
41	313	18.8	2660	12 ADN42341	Adn42341 Human CDN
42	313	18.8	3024	6 AAD38954	Aad38954 Human dip
43	313	18.8	3660	12 ADQ67504	Adq67504 Novel hum
44	313	18.8	3716	6 ABQ75955	Abq75955 Human PWM
45	313	18.8	4076	6 ABK83337	Abk83337 cDNA enco

ALIGNMENTS

RESULT 1
AAC85696
ID AAC85696 standard; cDNA; 1669 BP.

XX AAC85696;

DT 29-JUN-2001 (first entry)

DE Nucleotide sequence encoding human DPP8 318Thr-658Val+759Ala-882Ile.

XX Human; dipeptidyl aminopeptidase; DPP8; prollyl oligopeptidase;
KW dipeptidyl peptidase; DPPIV; T cell; cleavage; diarrhoea;
KW growth hormone deficiency; glucose level; mucosal regeneration;
KW non-insulin dependent diabetes mellitus; glucose intolerance;
KW immunosuppression; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

FT CDS 2..1399

FT /*tag= a

FT /product= "Human DPP8 318Thr-658Val+759Ala-882Ile"

XX WO200119866-A1.

XX 22-MAR-2001.

XX 11-SEP-2000; 2000WO-AU001085.

XX 10-SEP-1999; 99AU-00002762.

XX 18-FEB-2000; 2000AU-00005709.

XX (UNSY) UNIV SYDNEY.

XX Abbott CA, Gorell MD;

XX WPI; 2001-281520/29.

XX P-PSDB; AAB47189.

PT New human dipeptidyl aminopeptidase (DPP8) useful for cleaving
PT substrates, identifying inhibitors of DPP8 catalytic activity which have
PT therapeutic uses, and for detecting activated T cells.

XX Claim 21; Page 72-74; 78pp; English.

CC The sequences given in AAC85695-97 encode fragments of human dipeptidyl

CC aminopeptidase (DPP8). DPP8 has substrate specificity for H-Ala-Pro-pNA,

CC H-Gly-Pro-pNA and H-Arg-Pro-pNA. Therefore, it is a prolyl oligopeptidase

CC and a dipeptidyl peptidase, because it is capable of hydrolysing the

CC peptide bond C-terminal to Pro in each of these compounds. DPP8 is

CC homologous with human DPP4. DPP8 is useful for cleaving a substrate, and

CC for detecting an activated T cell which involves measuring the level of

CC DPP8 gene expression in a T cell. The level of DPP8 expression is

CC detected by detecting the amount of DPP8 RNA in the cell. It is also

CC useful for identifying a molecule capable of inhibiting the cleavage of

CC the substrate by DPP8. Molecules identified as inhibiting DPP8 catalytic

CC activity may be useful for treating diarrhoea, growth hormone deficiency,

CC lowering glucose levels in non-insulin dependent diabetes mellitus and

CC other disorders involving glucose intolerance, enhancing mucosal

CC regeneration and as immunosuppressants

XX SQ Sequence 1669 BP; 524 A; 336 C; 336 G; 473 T; 0 U; 0 Other;

Query Match 100.0%; Score 1669; DB 4; Length 1669;

Beat Local Similarity 100.0%; Pred. No. 0;

Matches 1669; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AACAGGTACAGCAAACTCTAAAGTCACCTTTTAAAGATGTCAGAAATATGATGCTGA 60

DB 1 AACAGGTACAGCAAACTCTAAAGTCACCTTTTAAAGATGTCAGAAATATGATGCTGA 60

QY 61 AGNAGGATCATAGATGTCATAGATAGGAACTATTCACCTTTTGAAGATCTATTTTGA 120

DB 61 AGNAGGATCATAGATGTCATAGATAGGAACTATTCACCTTTTGAAGATCTATTTTGA 120

QY 121 AGAGTTGAATATATTGCGAGAGCTGGATGGATCTCTGAGGAAATATGCTGGTCCAT 180

DB 121 AGAGTTGAATATATTGCGAGAGCTGGATGGATCTCTGAGGAAATATGCTGGTCCAT 180

QY 181 CCTACTAGATCGCTCCAGACCTGGCTACAGATAGTGTGTGATCTCACCTGAATTTTAT 240

DB 181 CCTACTAGATCGCTCCAGACCTGGCTACAGATAGTGTGTGATCTCACCTGAATTTTAT 240

QY 241 CCCAGTAGAAGATGATGTTATGGAAAGGCAGAGACTCAATTGAGTCAGTGGCTGATCTGT 300

DB 241 CCCAGTAGAAGATGATGTTATGGAAAGGCAGAGACTCAATTGAGTCAGTGGCTGATCTGT 300

QY 301 GAGGCCACTAATTTATCTATGAAGAAACACAGACATCTGGATAAATATCCATGACATCTT 360

DB 301 GAGGCCACTAATTTATCTATGAAGAAACACAGACATCTGGATAAATATCCATGACATCTT 360

QY 361 TCATGTTTTTCCCAAGTCACGAGAGGAAATTCAGTGTATTTTGGCTCTGAATGCAA 420

DB 361 TCATGTTTTTCCCAAGTCACGAGAGGAAATTCAGTGTATTTTGGCTCTGAATGCAA 420

QY 421 AACAGGTTTCCGTCATTTATACAAATATACATCTATTTTAAAGGAAACCAATATAACG 480

DB 421 AACAGGTTTCCGTCATTTATACAAATATACATCTATTTTAAAGGAAACCAATATAACG 480

QY 481 ATCCAGTGGGGCTCCGTCCTCCAGTGATTTTCAAGTGCTCTATCAAGGAGGATAGC 540

DB 481 ATCCAGTGGGGCTCCGTCCTCCAGTGATTTTCAAGTGCTCTATCAAGGAGGATAGC 540

QY 541 AATTACAGTGGTGAATGGAGTTCTTGGCCGGATGGATCTAATATCCAGTTGATGA 600

DB 541 AATTACAGTGGTGAATGGAGTTCTTGGCCGGATGGATCTAATATCCAGTTGATGA 600

QY 601 AGTCAGAGGCTGGTATATTTTGAAGGCACCAAGAGCTCCCTTTTAGAGCATACCTGTGA 660

DB 601 AGTCAGAGGCTGGTATATTTTGAAGGCACCAAGAGCTCCCTTTTAGAGCATACCTGTGA 660

QY 661 CGTAGTCAGTTACGTAATCTCTGGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACA 720

DB 661 CGTAGTCAGTTACGTAATCTCTGGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACA 720

QY 721 TTCTTGTGTCATCAGTCAGCAGCTGTGACCTCTTTTAAAGTAAGTATAGTACCAAGAGAA 780

DB 721 TTCTTGTGTCATCAGTCAGCAGCTGTGACCTCTTTTAAAGTAAGTATAGTACCAAGAGAA 780

QY 781 TCCACATGTCGTGTCCTTTTCAAGCTATCAAGTCTCAAGATGACCCCACTCTGCAAAAC 840

DB 781 TCCACATGTCGTGTCCTTTTCAAGCTATCAAGTCTCAAGATGACCCCACTCTGCAAAAC 840

QY 841 AAAGAAATTTTGGGCGCACCATTTTGGATTCAGCAGGTCCTCTTCTGACTATATCTCTCC 900

DB 841 AAAGAAATTTTGGGCGCACCATTTTGGATTCAGCAGGTCCTCTTCTGACTATATCTCTCC 900

QY 901 AGAAATTTTCTCTTTTGAAGTACTACTGATTTTACATGTATGATGGATGCTCTCAAGCC 960

DB 901 AGAAATTTTCTCTTTTGAAGTACTACTGATTTTACATGTATGATGGATGCTCTCAAGCC 960

QY 961 TCATGATCTACAGCCTGGAAAGAAATATCTTACTGCTGTTCATATATGATGGTGGCTCA 1020

DB 961 TCATGATCTACAGCCTGGAAAGAAATATCTTACTGCTGTTCATATATGATGGTGGCTCA 1020

QY 1021 GGTTCCTATTGCTGGGGCCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGA 1080

DB 1021 GGTTCCTATTGCTGGGGCCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGA 1080

QY 1081 ACGTTATATGGGTCACTGACCCAGAAATGAACAGGCTATTACTTAGGATCTGTGGCCAT 1140

DB 1081 ACGTTATATGGGTCACTGACCCAGAAATGAACAGGCTATTACTTAGGATCTGTGGCCAT 1140

QY 1141 GCAAGCAGAAAGTTCCTCTCTGAACCAATCGTTTACTGCTCTTACATGTTTCTCTGGA 1200

DB 1141 GCAAGCAGAAAGTTCCTCTCTGAACCAATCGTTTACTGCTCTTACATGTTTCTCTGGA 1200

QY 1201 TGAGAAATGTCATTTTGGCAGTATATTTAGTATTTTCTGAGTGGGCTGGGAA 1260

DB 1201 TGAGAAATGTCATTTTGGCAGTATATTTAGTATTTTCTGAGTGGGCTGGGAA 1260

QY 1261 GGCATATGATTTACAGATCTATCTCAGGAGACACAGCATAGAGTTCTTGAATCGGG 1320

DB 1261 GGCATATGATTTACAGATCTATCTCAGGAGACACAGCATAGAGTTCTTGAATCGGG 1320

QY 1321 AGAACATTTAGAACTGCTCTTTTGGCACTACCTTCAAGAAACCTTGGATCACGTAATTCG 1380

DB 1321 AGAACATTTAGAACTGCTCTTTTGGCACTACCTTCAAGAAACCTTGGATCACGTAATTCG 1380

QY 1381 TGCTCTAAAAAGTGATATAAATTTTGAACCTGTAGAACTCTCTGGTATACACTGGCTATTT 1440

DB 1381 TGCTCTAAAAAGTGATATAAATTTTGAACCTGTAGAACTCTCTGGTATACACTGGCTATTT 1440

QY 1441 AACCAAAATGAGGAGGTTTAAATCAACAGAAACACAGAAATTTGATCATCATTTTGTATACC 1500

DB 1441 AACCAAAATGAGGAGGTTTAAATCAACAGAAACACAGAAATTTGATCATCATTTTGTATACC 1500

QY 1501 TGCCATGTAACATCTACTCTGAAATAAATGTTGGTGGCCATGCGGGGTCTACGGTTTGT 1560

DB 1501 TGCCATGTAACATCTACTCTGAAATAAATGTTGGTGGCCATGCGGGGTCTACGGTTTGT 1560

QY 1561 GGTAGTAACTCTAATACCTTAAACCCCATGCTCAAAATCAAAATGATACATATCTCTGAGA 1620

DB 1561 GGTAGTAACTCTAATACCTTAAACCCCATGCTCAAAATCAAAATGATACATATCTCTGAGA 1620

QY 1621 GACCCAGCAATACCAATAGAAATTTACTTAAAAAABAAAAAABAAAAA 1669

DB 1621 GACCCAGCAATACCAATAGAAATTTACTTAAAAAABAAAAAABAAAAA 1669

RESULT 2

ABN59775

ID ABN59775 standard; cDNA; 2669 BP.

XX

AC ABN59775;

XX

DT 28-JUN-2002 (first entry)

XX

DE Novel human coding sequence SEQ ID NO: 186.
XX Human; antianaemic; vulnerary; antiinflammatory; immunomodulator;
KW antiinfertility; cerebroprotective; cytostatic; rheumatic; gene therapy;
KW neuroprotective; antiparkinsonian; protein therapy; EST;
KW expressed sequence tag; gene; ss.
XX Homo sapiens.
OS
XX WO200222660-A2.
XX
XX 21-MAR-2002.
XX
XX 10-SEP-2001; 2001WO-US026015.
XX
XX 11-SEP-2000; 2000US-00659671.
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
PI Xue AJ, Yang Y, Wehrman T, Drmanac RT;
PI
XX WPI: 2002-282408/33.
DR P-PSDB; ABB97362.
DR
XX An isolated polynucleotide for treating diseases associated with its
PT encoded polypeptide such as cancer and multiple sclerosis.
PT
XX Claim 1; SEQ ID NO 186; 509pp; English.
PS
XX The present invention provides the protein and coding sequences of 444
CC novel human proteins. These were isolated from expressed sequences tags
CC (ESTs). They can be used to stimulate cell growth, to regulate
CC haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth
CC e.g. in burh treatment, to regulate the immune system e.g. to treat
CC multiple sclerosis, to regulate activin or inhibin e.g. to treat
CC infertility, to regulate haemostasis or thrombolysis e.g. to treat stroke
CC and cancer, to screen for drugs, to treat inflammatory conditions e.g.
CC rheumatoid arthritis, and to treat nervous system disorders e.g.
CC Parkinson's disease. The present sequence is a coding sequence of the
CC invention
XX
SQ Sequence 2668 BP; 796 A; 564 C; 592 G; 716 T; 0 U; 0 Other;

Query Match 99.4%; Score 1659; DB 6; Length 2668;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1659; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AACAGGTACAGCAATCCCTAAAGTCACCTTTAAAGATGTCAGAAATAATGATGCTGA 60
DB 1010 AACAGGTACAGCAATCCCTAAAGTCACCTTTAAAGATGTCAGAAATAATGATGCTGA 1069

QY 61 AGAAGGATCATAGATGTCATAGATAAGGACTAATTCACCTTTTTCAGATTTCTATTGA 120
DB 1070 AGAAGGATCATAGATGTCATAGATAAGGACTAATTCACCTTTTTCAGATTTCTATTGA 1129

QY 121 AGGAGTTGAATATATTGGCAGAGCTGGATGGAATCCTGAGGAAATATGCTGGTCCAT 180
DB 1130 AGGAGTTGAATATATTGGCAGAGCTGGATGGAATCCTGAGGAAATATGCTGGTCCAT 1189

QY 181 CCTACTAGATCGCTCCAGACTCGCCTACAGATAGTGTGATCTCACCTGAATTTATTAT 240
DB 1190 CCTACTAGATCGCTCCAGACTCGCCTACAGATAGTGTGATCTCACCTGAATTTATTAT 1249

QY 241 CCAGTAGAAGATGATGTTATGGAAAGGAGAGACTCATTTGAGTCAGTGGCTGATTCGT 300
DB 1250 CCAGTAGAAGATGATGTTATGGAAAGGAGAGACTCATTTGAGTCAGTGGCTGATTCGT 1309

QY 301 GAGCCACATAATTTCTATGAAGAAACAACAGACATCTGGATAAATATCCATGACATCTT 360
DB 1310 GAGCCACATAATTTCTATGAAGAAACAACAGACATCTGGATAAATATCCATGACATCTT 1369

QY 361 TCATGTTTTTCCCCAAGTCACGAAGAGGAAATTGAGTTTATTTTTTGCCTCTGAATGCAA 420

DB 1370 TCATGTTTTTCCCCAAGTCACGAAGAGAAATTGAGTTTATTTTGGTCTCTGAATGCAA 1429
QY 421 AACAGGTTTCCGTCATTATACAAAATTACATCTATTTTAAAGGAAAGCAAAATATAAACG 480
DB 1430 AACAGGTTTCCGTCATTATACAAAATTACATCTATTTTAAAGGAAAGCAAAATATAAACG 1489
QY 481 ATCCAGTGGTGGCTGCTGCTCCAAAGTGATTTCAAGTGTCTATCAAAAGAGGAGATAGC 540
DB 1490 ATCCAGTGGTGGCTGCTGCTCCAAAGTGATTTCAAGTGTCTATCAAAAGAGGAGATAGC 1549
QY 541 AATTACCAGTGGTGAATGGGAAGTCTTGGCCGGCATGGATCTAATATCCCAAGTTGATGA 600
DB 1550 AATTACCAGTGGTGAATGGGAAGTCTTGGCCGGCATGGATCTAATATCCCAAGTTGATGA 1609
QY 601 AGTCAGAAAGCTGGTATATTTTGAAGGCCAACAAAGACTCCCTTTTAGAGCATCACCTGTA 660
DB 1610 AGTCAGAAAGCTGGTATATTTTGAAGGCCAACAAAGACTCCCTTTTAGAGCATCACCTGTA 1669
QY 661 CGTAGTCAGTTAGCTAAATCCCTGGAGAGGTGACAGGCTGACGACCGTGGCTACTACAA 720
DB 1670 CGTAGTCAGTTAGCTAAATCCCTGGAGAGGTGACAGGCTGACGACCGTGGCTACTACAA 1729
QY 721 TTCTTGCTGCATCAGTCAGCAGCTGTGACTTCTTTTAAAGTAAAGTATAGTAAACCAAGAA 780
DB 1730 TTCTTGCTGCATCAGTCAGCAGCTGTGACTTCTTTTAAAGTAAAGTATAGTAAACCAAGAA 1789
QY 781 TCCACACTGTGTCTCCCTTTTACAAGCTATCAAGTCTCTGAAGATGACCCAACTTGCANAAAC 840
DB 1790 TCCACACTGTGTCTCCCTTTTACAAGCTATCAAGTCTCTGAAGATGACCCAACTTGCANAAAC 1849
QY 841 AAAGGAATTTTGGGCCACCAATTTGGATTCAGCAGGCTCTCTTCCTGACTATACCTCTCC 900
DB 1850 AAAGGAATTTTGGGCCACCAATTTGGATTCAGCAGGCTCTCTTCCTGACTATACCTCTCC 1909
QY 901 AGAAATTTTCTCTTTTGAAGTACTACTGATTTACATTTGATGGATGCTCTACAAGCC 960
DB 1910 AGAAATTTTCTCTTTTGAAGTACTACTGATTTACATTTGATGGATGCTCTACAAGCC 1969
QY 961 TCATGATCTACAGCTGGAAAGAAATATCTACTGTCTGTGTTTATATATGTTGGTCTCTCA 1020
DB 1970 TCATGATCTACAGCTGGAAAGAAATATCTACTGTCTGTGTTTATATGTTGGTCTCTCA 2029
QY 1021 GGTGTCTATTTGGGGCCCCAGTCTACTGTGGATCTTCTATGATACAGGATACACGGA 1080
DB 2030 GGTGTCTATTTGGTGGGGCCCCAGTCTACTGTGGATCTTCTATGATACAGGATACACGGA 2089
QY 1081 ACGTTATATGGGTCAACCTGACCCAGAAATGAACAGGCTATTTACTTAGGATCTGTGGCCAT 1140
DB 2090 ACGTTATATGGGTCAACCTGACCCAGAAATGAACAGGCTATTTACTTAGGATCTGTGGCCAT 2149
QY 1141 GCAAGCAGAAAAGTTCCTCTCTGAACCAATCGTTTACTGCTCTTACATGGTTTCTCTGGA 1200
DB 2150 GCAAGCAGAAAAGTTCCTCTCTGAACCAATCGTTTACTGCTCTTACATGGTTTCTCTGGA 2209
QY 1201 TGAGAAATGTCATTTTGGCACAATACAGATATATTTACTGAGTGTGTTTGTAGGGCTGGA 1260
DB 2210 TGAGAAATGTCATTTTGGCACAATACAGATATATTTACTGAGTGTGTTTGTAGGGCTGGA 2269
QY 1261 GCCATATGATTTACAGATCTATCTCAGAGAGACACAGATAGAGTTTCTGATTCGGG 1320
DB 2270 GCCATATGATTTACAGATCTATCTCAGAGAGACACAGATAGAGTTTCTGATTCGGG 2329
QY 1321 AGACATTTATGAATGTCATCTTTTGCACCTACCTTCAAGAAAACCTTGGATCACGTTATTCG 1380
DB 2330 AGACATTTATGAATGTCATCTTTTGCACCTACCTTCAAGAAAACCTTGGATCACGTTATTCG 2389
QY 1381 TGCTCTAAAAGTGATATAATTTTACCTGTGTAGAACTCTCTGGTATACACTGGCTATTTT 1440
DB 2390 TGCTCTAAAAGTGATATAATTTTGAACCTGTGTAGAACTCTCTGGTATACACTGGCTATTTT 2449
QY 1441 AACCAATATGAGGAGTTTAAATCAACAGAAAACACAGAAATTCATCATCATTTTGAATACC 1500

QY 1141 GCAGCAGAAAAGTTCCCTCTGAAACCAAAATCGTTTACTGCTCTTACATGAGTTCCTCGGA 1200
DB 2324 GCAAGCAGAAAAGTTCCCTCTGAAACCAAAATCGTTTACTGCTCTTACATGAGTTCCTCGGA 2383
QY 1201 TGAGAGTCTCCATTTTCACATACCACTATATTACTGAGTCTTTTGTAGTGGGCTGGAAA 1260
DB 2384 TGAGAGTCTCCATTTTCACATACCACTATATTACTGAGTCTTTTGTAGTGGGCTGGAAA 2443
QY 1261 GCCATATGATTTACAGATCTATCTCCTCAGGAGACACAGCATAGAGTTCCTGAATCGGG 1320
DB 2444 GCCATATGATTTACAGATCTATCTCCTCAGGAGACACAGCATAGAGTTCCTGAATCGGG 2503
QY 1321 AGACATATGAACTGATCTTTTGCATCTACCTTCAAGAAAACCTTGGATCACTGATTCG 1380
DB 2504 AGACATATGAACTGATCTTTTGCATCTACCTTCAAGAAAACCTTGGATCACTGATTCG 2563
QY 1381 TGCTCTAAAGTGATATAATTTTGCCTGCTGTAGACTCTCTGCTATACACTGGCTATTT 1440
DB 2564 TGCTCTAAAGTGATATAATTTTGCCTGCTGTAGACTCTCTGCTATACACTGGCTATTT 2623
QY 1441 AACCAAAATGAGAGGTTTAAATCAACAGAAAACACAGAAATGATCATCATATTTGATACC 1500
DB 2624 AACCAAAATGAGAGGTTTAAATCAACAGAAAACACAGAAATGATCATCATATTTGATACC 2683
QY 1501 TGCCATGTAACATCTACTCTCTGAAAATAAATGTTGGTGCATCCAGGGGTCTACGGTTTGT 1560
DB 2684 TGCCATGTAACATCTACTCTCTGAAAATAAATGTTGGTGCATCCAGGGGTCTACGGTTTGT 2743
QY 1561 GGTAGTAACTTAATACCTTAACCCCAATGCTCAAAATCAAAATGATATATTTCTTGAGA 1620
DB 2744 GGTAGTAACTTAATACCTTAACCCCAATGCTCAAAATCAAAATGATATATTTCTTGAGA 2803
QY 1621 GACCCAGCAATACCAATGAATTAATCAAAATCAAAATCAAAATCAAAATCAAAATCAAA 1659
DB 2804 GACCCAGCAATACCAATGAATTAATCAAAATCAAAATCAAAATCAAAATCAAAATCAAA 2842

RESULT 4

ID ABK83325 standard; cDNA; 4523 BP.
XX ABK83325;

AC ABK83325;
DT 12-AUG-2002 (first entry)
DE cDNA encoding human DPRP-1 splice variant #1.

KW Human; serine protease; dipeptidyl peptidase IV-related protein; DPRP;
KW DPPIV; infection; human immunodeficiency virus; HIV-1; HIV-2; pain;
KW diabetes; infertility; obesity; anorexia; Parkinson's disease; stroke;
KW heart failure; hypertension; urinary retention; osteoporosis; cancer;
KW ulcer; allergy; cancer; psychotic disorder; neurological disorder;
KW dyskinesia; reproductive disorder; inflammatory disorder;
KW metabolic disorder; gene; ss.

XX Homo sapiens.
OS
XX WO200231134-A2.

PN
XX
XX PD 18-APR-2002.
XX PF 12-OCT-2001; 2001WO-US031874.

PR 12-OCT-2000; 2000US-0240117P.
XX
XX PA (FERR) FERRING BV.
XX

XX Qi S, Akinsanya KO, Riviere PJ, Junien J;
PI WPI; 2002-444178/47.
XX P-PSDB; ABG61594.
DR
XX

PT New dipeptidyl peptidase IV-related proteins and nucleic acids encoding
PT the proteins, useful for treating e.g. fungal, bacterial, protozoan and
PT viral infections, cancers, allergies, neurological disorders, or pain.
XX Disclosure; Page 61-62; 113pp; English.

XX The present invention relates to the isolation of novel human serine
CC proteases referred to as dipeptidyl peptidase IV (DPPIV)-related proteins
CC (DPRP). The dipeptidyl peptidase IV-related proteins (DPRP) and nucleic
CC acids encoding them are useful for treating infections such as fungal,
CC bacterial, protozoan and viral infections, particularly infections caused
CC by human immunodeficiency virus (HIV-1 or HIV-2), pain, diabetes,
CC precocious puberty, infertility, obesity, anorexia, bulimia, Parkinson's
CC disease, acute heart failure, hypotension, hypertension, urinary
CC retention, osteoporosis, angina pectoris, stroke, ulcers, asthma,
CC allergies, cancers, migraine, vomiting, or psychotic and neurokinetic
CC disorders (e.g. anxiety, dementia, or schizophrenia), and dyskinesias.
CC These may also be used in discovering therapeutic agents for the
CC treatment of reproductive, inflammatory and metabolic disorders. ABK83322
CC -ABK83343 encode human DPRP proteins
XX

QY Sequence 4523 BP; 1384 A; 828 C; 940 G; 1371 T; 0 U; 0 Other;
XX

Query Match 98.8%; Score 1648.4; DB 6; Length 4523;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1663; Conservative 0; Mismatches 1; Indels 4; Gaps 1;

QY 1 AACAGGTACACCAATCTTAAAGTCACTTTTAAAGATGTCAGAAATATGATTGATGCTGA 60
DB 1164 AACAGGTACACCAATCTTAAAGTCACTTTTAAAGATGTCAGAAATATGATTGATGCTGA 1223
QY 61 AGGAAGGATCATAGATGTCATAGATAAGGAACATAATCAACCTTTTGGATTTCTATTGA 120
DB 1224 AGGAAGGATCATAGATGTCATAGATAAGGAACATAATCAACCTTTTGGATTTCTATTGA 1283
QY 121 AGGAGTTGAATATATTCGAGAGCTGGATGGACTCTGAGGGAATAATGCTTGGTCCAT 180
DB 1284 AGGAGTTGAATATATTCGAGAGCTGGATGGACTCTGAGGGAATAATGCTTGGTCCAT 1343
QY 181 CCTACTAGATCGCTCCAGACTCGCTTACAGATAGTGTGATCTCACCTGAAATTTTAT 240
DB 1344 CCTACTAGATCGCTCCAGACTCGCTTACAGATAGTGTGATCTCACCTGAAATTTTAT 1403
QY 241 CCCAGTAGAAGATGATGTTATGGAAGGAGAGACTCATTGAGTCAGTGCCTGATTTCTGT 300
DB 1404 CCCAGTAGAAGATGATGTTATGGAAGGAGAGACTCATTGAGTCAGTGCCTGATTTCTGT 1463
QY 301 GACGCCACTAATTTATCTATGAAGAAACACAGACATCTCGGATAAATATCCATGACATCTT 360
DB 1464 GACGCCACTAATTTATCTATGAAGAAACACAGACATCTCGGATAAATATCCATGACATCTT 1523
QY 361 TCATGTTTTTCCCAAGTCCAGAGGAAATTTGAGTTTATTTTGGCTCTGAAATGCAA 420
DB 1524 TCATGTTTTTCCCAAGTCCAGAGGAAATTTGAGTTTATTTTGGCTCTGAAATGCAA 1583
QY 421 AACAGGTTTCCGTCATTTATACAAATTTACATCTATTTTAAAGGAAACAAATATAAAGC 480
DB 1584 AACAGGTTTCCGTCATTTATACAAATTTACATCTATTTTAAAGGAAACAAATATAAAGC 1643
QY 481 ATCCAGTGGTGGGCTGCTGCTCCAAAGTGAATTTCAAGTGTCTTATCAAGAGGAGATAGC 540
DB 1644 ATCCAGTGGTGGGCTGCTGCTCCAAAGTGAATTTCAAGTGTCTTATCAAGAGGAGATAGC 1703
QY 541 AATTACCAGTGGTGAATGGGAAGTTCTTGGCCGGCATGGATCTAATATCCAGTTGATGA 600
DB 1704 AATTACCAGTGGTGAATGGGAAGTTCTTGGCCGGCATGGATCTAATATCCAGTTGATGA 1763
QY 601 AGTCAGAGGCTGGTATATTTTGAAGGCCAACAAAGACTCCCTTTTAGAGCATCACCTGTA 660
DB 1764 AGTCAGAGGCTGGTATATTTTGAAGGCCAACAAAGACTCCCTTTTAGAGCATCACCTGTA 1823
QY 661 CGTAGTCAGTTACGTAAATCCTCGAGAGGTGACAAAGGCTGACTGACCGTGGCTACTCACA 720
|||||

Db	1824	CGTAGTCAGTTACGTAATAATCCCTGGAGAGGTCGACAAAGGCTGACTGACCGTGGCTACTCACA	1883
Qy	721	TTCTTCTGCTGATCAGTCAGCAGCTGTGACTTCTTTTAAAGTAAGTATAGTAACAGAA	780
Db	1884	TTCTTCTGCTGATCAGTCAGCAGCTGTGACTTCTTTTAAAGTAAGTATAGTAACAGAA	1943
Qy	781	TCCACACTGTGTGCTCCCTTTTACAAGCTATCAAGTCTGAAGTACCCAACTTGC AAAAC	840
Db	1944	TCCACACTGTGTGCTCCCTTTTACAAGCTATCAAGTCTGAAGTACCCAACTTGC AAAAC	2003
Qy	841	AAAGGAATTTTGGGCGCACCATTTTGGATTTCAAGCTCTCTTCTGACTATATCTCTCC	900
Db	2004	AAAGGAATTTTGGGCGCACCATTTTGGATTTCAAGCTCTCTTCTGACTATATCTCTCC	2063
Qy	901	AGAAATTTTCTCTTTTGAAGTACTACTGGATTTACATTTGTATGGGATGCTCTACAAGCC	960
Db	2064	AGAAATTTTCTCTTTTGAAGTACTACTGGATTTACATTTGTATGGGATGCTCTACAAGCC	2123
Qy	961	TCATGATCTACAGCTTGAAGAAATATCTTACTGTGCTGTTTATATATGTTGTTCTCTCA	1020
Db	2124	TCATGATCTACAGCTTGAAGAAATATCTTACTGTGCTGTTTATATATGTTGTTCTCTCA	2179
Qy	1021	GGTTGCTATTGCTGGGGCCCCAGTCACCTCTGTGGATCTTCTATGATACAGGATACACGGA	1080
Db	2180	GGTTGCTATTGCTGGGGCCCCAGTCACCTCTGTGGATCTTCTATGATACAGGATACACGGA	2239
Qy	1081	ACGTTATATGGGTCACCTCTGACAGAAATGAACAGGCTTACTTAGGATCTCTGGCCAT	1140
Db	2240	ACGTTATATGGGTCACCTCTGACAGAAATGAACAGGCTTACTTAGGATCTCTGGCCAT	2299
Qy	1141	GCAAGCAGAAAGTCCCTCTGAACCAATCGTTTACTGCTCTTACATGTTTCTCTGGA	1200
Db	2300	GCAAGCAGAAAGTCCCTCTGAACCAATCGTTTACTGCTCTTACATGTTTCTCTGGA	2359
Qy	1201	TGAGAAATGTCATTTTGCAATACACAGTATATTTACTGAGTTTTTTTGTAGGGGCTGAAA	1260
Db	2360	TGAGAAATGTCATTTTGCAATACACAGTATATTTACTGAGTTTTTTTGTAGGGGCTGAAA	2419
Qy	1261	GCCATATGATTTTACAGATCTATCCTCAGAGAGACACAGCATATAAGTTCTGAAATCGGG	1320
Db	2420	GCCATATGATTTTACAGATCTATCCTCAGAGAGACACAGCATATAAGTTCTGAAATCGGG	2479
Qy	1321	AGAACATTTATGAACTGCTCTTTTGCACTACCTTCAAGAAAACCTTGGATCAGGTATTC	1380
Db	2480	AGAACATTTATGAACTGCTCTTTTGCACTACCTTCAAGAAAACCTTGGATCAGGTATTC	2539
Qy	1381	TGCTCTAAAAGTGATATAATTTTGACCTGTGTAGAACTCTCTGGTATACACTGGCTATTT	1440
Db	2540	TGCTCTAAAAGTGATATAATTTTGACCTGTGTAGAACTCTCTGGTATACACTGGCTATTT	2599
Qy	1441	AACCAATGAGGAGGTTTAAATCAACAGAAAACACAGAAATTCATCATCATTTTGTATACC	1500
Db	2600	AACCAATGAGGAGGTTTAAATCAACAGAAAACACAGAAATTCATCATCATTTTGTATACC	2659
Qy	1501	TGCCATGTAACATCTACTCTGAAATAAATGTGTGTCATGCGGGGTCTACGGTTTGT	1560
Db	2660	TGCCATGTAACATCTACTCTGAAATAAATGTGTGTCATGCGGGGTCTACGGTTTGT	2719
Qy	1561	GGTAGTAATCTATATCTTAAACCCACATGCTCAAAATCAATGATACATATTCCTTGAGA	1620
Db	2720	GGTAGTAATCTATATCTTAAACCCACATGCTCAAAATCAATGATACATATTCCTTGAGA	2779
Qy	1621	GACCCAGCATACCATTAAGATTTACTTAAAAAATAAAAAAATAAAAAA 1668	
Db	2780	GACCCAGCATACCATTAAGATTTACTTAAAAAATAAAAAAATAAAAAA 2827	
RESULT 5			
ADT04072			
XX	ADT04072 standard; DNA; 2702 BP.		
AC	ADT04072;		
XX			

DT

30-DEC-2004 (first entry)

Human protein modification and maintenance molecule gene SeqID52.

protein modification and maintenance molecule; PMMW; cytostatic; immunomodulator; expression; immune disorder; cancer; gene; ds; human.

Homo sapiens.

Key Location/Qualifiers

CDS /tag= a

FT /product= "Human protein modification and maintenance molecule"

WO2004084806-A2.

07-OCT-2004.

16-MAR-2004; 2004WO-US0080006.

21-MAR-2003; 2003US-0456864P.

03-APR-2003; 2003US-0460512P.

19-MAY-2003; 2003US-0472027P.

30-MAY-2003; 2003US-0475072P.

30-MAY-2003; 2003US-0475190P.

24-JUN-2003; 2003US-0482575P.

13-AUG-2003; 2003US-0495151P.

(INCY-) INCYTE CORP.

Baughn MR, Marquis JP, Kable AE, Chawla NK, Emerling BM, Lee SY; Hafalia AJA, Ramkumar J, Richardson TW, Wang JT, Nakamura LL; Yang J, Jin P, Becha SB, Wilson AD;

WPI; 2004-710256/69.

P-PSDB; ADT04044.

New human protein modification and maintenance molecules (PMMW) polypeptide, useful in preparing a composition for treating a disease PT associated with decreased expression or overexpression of PMMW e.g., cancer.

Claim 5; SEQ ID NO 52; 222pp; English.

This invention relates to novel human protein modification and maintenance molecule (PMMW) proteins and the DNA sequences which encode them. The invention may be useful for the production of compounds with a cytostatic or immunomodulator activity. The proteins are useful in preparing a composition for diagnosing or treating a disease or condition associated with decreased expression or overexpression of functional PMMW, for example immune disorders or cancer. The present sequence is that of a gene which encodes a human protein modification and maintenance molecule (PMMW) protein of the invention.

Sequence 2702 BP; 798 A; 569 C; 609 G; 726 T; 0 U; 0 Other;

Query Match 96.7%; Score 1614.4; DB 13; Length 2702;

Best Local Similarity 99.2%; Pred. No. 0;

Matches 1634; Conservative 0; Mismatches 11; Indels 3; Gaps 1;

Qy 1 AACAGGTACAGCAAACTCTAAAGTCACCTTTTAAGATGTCAGAAAATAATGATTGCTGGA 60

Db 1003 AACAGGTACAGCAAACTCTAAAGTCACCTTTTAAGATGTCAGAAAATAATGATTGCTGGA 1062

Qy 61 AGGAAGGATCATAGATGTCATAGATAGGAACATAATCAACCTTTTGAGATTTCTATTGA 120

Db 1063 AGGAAGGATCATAGATGTCATAGATAGGAACATAATCAACCTTTTGAGATTTCTATTGA 1122

Qy 121 AGGAAGGATCATAGATGTCATAGATAGGAACATAATCAACCTTTTGAGATTTCTATTGA 180

Db 1123 AGGAAGGATCATAGATGTCATAGATAGGAACATAATCAACCTTTTGAGATTTCTATTGA 1182

QY 181 CCTACTAGATCGCTCCAGACTCGCCTACAGATAGTGTGATCTCACCTGAATTAATTTAT 240
DB 1183 CCTACTAGATCGCTCCAGACTCGCCTACAGATAGTGTGATCTCACCTGAATTAATTTAT 1242
QY 241 CCCAGTGAAGATGATGTTATGAAAGGAGAGAGACTCATGAGTCAGTGCCTGATTCGT 300
DB 1243 CCCAGTGAAGATGATGTTATGAAAGGAGAGAGACTCATGAGTCAGTGCCTGATTCGT 1302
QY 301 GACGCCACTAATTTATCTATGAAGAAACAAACAGACATCTGGATAAATATCCATGACATCTT 360
DB 1303 GACGCCACTAATTTATCTATGAAGAAACAAACAGACATCTGGATAAATATCCATGACATCTT 1362
QY 361 TCATGTTTTTCCCAAGTCAAGAGGAAATGAGTTATTTTGGCTCTGATGCAA 420
DB 1363 TCATGTTTTTCCCAAGTCAAGAGGAAATGAGTTATTTTGGCTCTGATGCAA 1422
QY 421 AACAGGTTTCGGTCAATTTATACAAATTAATCTATTTTAAAGGAAACAAATATAACG 480
DB 1423 AACAGGTTTCGGTCAATTTATACAAATTAATCTATTTTAAAGGAAACAAATATAACG 1482
QY 481 ATCCAGTGTGGGCTGCTGCTCCAAAGTGAATTTCAAGTGTCTCTATCAAGAGGAGATAGC 540
DB 1483 ATCCAGTGTGGGCTGCTGCTCCAAAGTGAATTTCAAGTGTCTCTATCAAGAGGAGATAGC 1542
QY 541 AATTACAGTGTGGAATGGGAAGTTCTTGGCCGGCATGGAATCTAATATCCAAAGTTGATGA 600
DB 1543 AATTACAGTGTGGAATGGGAAGTTCTTGGCCGGCATGGAATCTAATATCCAAAGTTGATGA 1602
QY 601 AGTCAGAGGCTGGTATATTTTGAAGGCACCAAGACTCCCTTTAGAGCATACCTGTA 660
DB 1603 AGTCAGAGGCTGGTATATTTTGAAGGCACCAAGACTCCCTTTAGAGCATACCTGTA 1662
QY 661 CGTAGTCAGTTACGTAAATCCTGGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACA 720
DB 1663 CGTAGTCAGTTACGTAAATCCTGGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACA 1722
QY 721 TTCTTGCTGTCAGTCAGTCAGCTGTGACTTTCTTTAATAAGTAAAGTATAGTAAACCAAGAGAA 780
DB 1723 TTCTTGCTGTCAGTCAGTCAGCTGTGACTTTCTTTAATAAGTAAAGTATAGTAAACCAAGAGAA 1782
QY 781 TCCACACTGTGTCTCCCTTTACAGCTATCAAGTCTCAAGTCTCAAGTACCCCACTTCCAAAC 840
DB 1783 TCCACACTGTGTCTCCCTTTACAGCTATCAAGTCTCAAGTCTCAAGTACCCCACTT---GCANA 1839
QY 841 AAGGAATTTTGGGCGACACATTTTGGATTTCAGCAGTCTCTTCTGACTATATCTCTCC 900
DB 1840 AAGGAATTTTGGGCGACACATTTTGGATTTCAGCAGTCTCTTCTGACTATATCTCTCC 1899
QY 901 AGAAATTTCTTTTTGAAAGTACTACTGGAATTTACATGTATGGATGCTCTTACAAGCC 960
DB 1900 AGAAATTTCTTTTTGAAAGTACTACTGGAATTTACATGTATGGATGCTCTTACAAGCC 1959
QY 961 TCATGATCTACAGCTGGAAGAAATATCCTACTGCTGTGTTTCATATATGTTGCTCTCA 1020
DB 1960 TCATGATCTACAGCTGGAAGAAATATCCTACTGCTGTGTTTCATATATGTTGCTCTCA 2019
QY 1021 GGTGTATTTGCTGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGATACACGGA 1080
DB 2020 GGTGTATTTGCTGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGATACACGGA 2079
QY 1081 ACGTATATATGGGTCAACCTGACAGAAATGAACAGGGCTATTACTTAGGATCTGTGGCCAT 1140
DB 2080 ACGTTATATGGGTCAACCTGACAGAAATGAACAGGGCTATTACTTAGGATCTGTGGCCAT 2139
QY 1141 GCAAGCAGAAAAGTTTCCCTCTGAACCAAATCGTTTACTGCTTTACATGTTTCCCTGGA 1200
DB 2140 GCAAGCAGAAAAGTTTCCCTCTGAACCAAATCGTTTACTGCTTTACATGTTTCCCTGGA 2199
QY 1201 TGAGAATGTCATTTTGGCATACACCATGATATTTACTGAGTTTTTTAGTGGGCTGGAAA 1260
DB 2200 TGAGAATGTCATTTTGGCATACACCATGATATTTACTGAGTTTTTTAGTGGGCTGGAAA 2259
QY 1261 GCCATATGATTTACAGATCTATCTCCTCAGGAGAGACACAGCATTAAGAGTTCTCTGAATCGGG 1320

DB 2260 GCCATATGATTTACAGATCTATCTCCTCAGGAGACACAGCATTAAGAGTTCTCTGAATCGGG 2319
QY 1321 AGAACATTTATGAACCTGCATCTTTTGGCACTACCTTCAAGAAAAACCTTGGATCACGTATTGC 1380
DB 2320 AGAACATTTATGAACCTGCATCTTTTGGCACTACCTTCAAGAAAAACCTTGGATCACGTATTGC 2379
QY 1381 TGCTCTAAAAAGTGATATAAATTTTGGACCTGTGTAGAACTCTCTGGTATACACCTGGCTATTT 1440
DB 2380 TGCTCTAAAAAGTGATATAAATTTTGGACCTGTGTAGAACTCTCTGGTATACACCTGGCTATTT 2439
QY 1441 AACCAATGAGGAGGTTTAAATCAACAGAAAAACACAGAAATTTGATCATCATTTTGTATACC 1500
DB 2440 AACCAATGAGGAGGTTTAAATCAACAGAAAAACACAGAAATTTGATCATCATTTTGTATACC 2499
QY 1501 TGCCATGTAAACATCTACTCTCTGAAATAAATGTGGTGGCCATGCAGGGGTCTACGGTTTGT 1560
DB 2500 TGCCATGTAAACATCTACTCTCTGAAATAAATGTGGTGGCCATGCAGGGGTCTACGGTTTGT 2559
QY 1561 GGTAGTAAATCTAATACCTTAAACCCCATGCTCAAAATCAAAATGATACATATTTCTCTGAGA 1620
DB 2560 GGTAGTAAATCTAATACCTTAAACCCCATGCTCAAAATCAAAATGATACATATTTCTCTGAGA 2619
QY 1621 GACCCAGCAATACCATTAAGAAATTTACTAA 1648
DB 2620 GACCCAGCAATACCATTAAGAAATTTACTAA 2647

RESULT 6
ABK83331
ID ABK83331 standard; cDNA; 4676 BP.
XX
AC ABK83331;
XX
DT 12-AUG-2002 (first entry)
XX
DE cDNA encoding human DPRP-1 splice variant #7.
KW Human; serine protease; dipeptidyl peptidase IV-related protein; DPRP;
KW DPPIV; infection; human immunodeficiency virus; HIV-1; HIV-2; pain;
KW diabetes; infertility; obesity; anorexia; Parkinson's disease; stroke;
KW heart failure; hypertension; urinary retention; osteoporosis; cancer;
KW ulcer; allergy; cancer; psychotic disorder; neurological disorder;
KW dyskenesia; reproductive disorder; inflammatory disorder;
KW metabolic disorder; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200231134-A2.
XX
PD 18-APR-2002.
XX
PF 12-OCT-2001; 2001WO-US031874.
XX
PR 12-OCT-2000; 2000US-0240117P.
XX
XX (FERR) FERRING BV.
XX
PI Qi S, Akinsanya KO, Riviere PJ, Junien J;
XX
DR WPI; 2002-444178/47.
DR P-PSDB; ABG61600.
XX
PT New dipeptidyl peptidase IV-related proteins and nucleic acids encoding
PT the proteins, useful for treating e.g. fungal, bacterial, protozoan and
PT viral infections, cancers, allergies, neurological disorders, or pain.
XX
PS Disclosure; Page 72-73; 113pp; English.
XX
CC The present invention relates to the isolation of novel human serine
CC proteases referred to as dipeptidyl peptidase IV (DPPIV)-related proteins
CC (DPRP). The dipeptidyl peptidase IV-related proteins (DPRP) and nucleic
CC acids encoding them are useful for treating infections such as fungal,

CC bacterial, protozoan and viral infections, particularly infections caused
CC by human immunodeficiency virus (HIV-1 or HIV-2), pain, diabetes,
CC precocious puberty, infertility, obesity, anorexia, bulimia, Parkinson's
CC disease, acute heart failure, hypotension, hypertension, urinary
CC retention, osteoporosis, angina pectoris, stroke, ulcers, asthma,
CC allergies, cancers, migraine, vomiting, psychotic and neurological
CC disorders (e.g. anxiety, dementia, or schizophrenia), and dyskinetic
CC These may also be used in discovering therapeutic agents for the
CC treatment of reproductive, inflammatory and metabolic disorders. ABK83322
CC -ABK83343 encode human DRP proteins
XX
SQ Sequence 4676 BP; 1424 A; 859 C; 979 G; 1414 T; 0 U; 0 Other;

Query Match 89.9%; Score 1501; DB 6; Length 4676;
Best Local Similarity 91.5%; Pred. No. 0;
Matches 1663; Conservative 5; Mismatches 149; Gaps 1;

QY 1 AACAGGTACAGCAATCTTAAGTCACCTTTAAAGTGTCAAGAAATATGATCTCATCGA 60
DB 1164 AACAGGTACAGCAATCTTAAGTCACCTTTAAAGTGTCAAGAAATATGATCTCATCGA 1223

QY 61 AGAAGGATCATAGATGTCATAGATAGCAATTAAGCACTTTTGAAGTCTTATTTGA 120
DB 1224 AGAAGGATCATAGATGTCATAGATAGCAATTAAGCACTTTTGAAGTCTTATTTGA 1283

QY 121 AGAGTGAATATATGCGAGAGCTGGATGCACTCTGAGGAAATATGCTTGGTCCAT 180
DB 1284 AGAGTGAATATATGCGAGAGCTGGATGCACTCTGAGGAAATATGCTTGGTCCAT 1343

QY 181 CCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTACCTGAATATTTAT 240
DB 1344 CCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTACCTGAATATTTAT 1403

QY 241 CCCAGTAGAAGTATGTTATGAAAGGACAGAGACTCATTTGAGTGTGCTGATCTGT 300
DB 1404 CCCAGTAGAAGTATGTTATGAAAGGACAGAGACTCATTTGAGTGTGCTGATCTGT 1463

QY 301 GAGCCACTAATATCTATGAAGAAACACAGACATCTGGATAATATCCATGACATCTT 360
DB 1464 GAGCCACTAATATCTATGAAGAAACACAGACATCTGGATAATATCCATGACATCTT 1523

QY 361 TCATGTTTTTCCCAAGTACAGAGAGAAATGAGTTTATTTTTGGCTCTGAATGCAA 420
DB 1524 TCATGTTTTTCCCAAGTACAGAGAGAAATGAGTTTATTTTTGGCTCTGAATGCAA 1583

QY 421 AACAGGTTCCGTCATTTATACAAATTTACATCTATTTAAAGGAAAGCAATATATAACG 480
DB 1584 AACAGGTTCCGTCATTTATACAAATTTACATCTATTTAAAGGAAAGCAATATATAACG 1643

QY 481 ATCCAGTGTGGGCTGCTGCTCCAGTGATTTCAAGTGTCTATCAAGAGGAGATAGC 540
DB 1644 ATCCAGTGTGGGCTGCTGCTCCAGTGATTTCAAGTGTCTATCAAGAGGAGATAGC 1703

QY 541 AATTACAGTGTGATGGAATGGAATGTTGGCGGCGATGGATCTAATATCCAAGTGTGATGA 600
DB 1704 AATTACAGTGTGATGGAATGGAATGTTGGCGGCGATGGATCTAATATCCAAGTGTGATGA 1763

QY 601 AGTCAGAAGGCTGGTATATTTTGAAGGACCAAGACCTCCCTTTAGAGCATCACCTGTA 660
DB 1764 AGTCAGAAGGCTGGTATATTTTGAAGGACCAAGACCTCCCTTTAGAGCATCACCTGTA 1823

QY 661 CGTAGTCAGTTACGTAATATCTGGAGAGGTGACAAAGCTGACTGACCGTGGCTACTACA 720
DB 1824 CGTAGTCAGTTACGTAATATCTGGAGAGGTGACAAAGCTGACTGACCGTGGCTACTACA 1883

QY 721 TTCTTCTGATCAGTCAGCTGACTGACTTCTTTTAAAGTAAGTATAGTAACCAAGAGAA 780
DB 1884 TTCTTCTGATCAGTCAGCTGACTGACTTCTTTTAAAGTAAGTATAGTAACCAAGAGAA 1943

QY 781 TCCACACTGTGTGCTCCCTTTTACAAGCTATCAAGTCTGAAAGATGACCCAACTTGC AAAAC 840
DB 1944 TCCACACTGTGTGCTCCCTTTTACAAGCTATCAAGTCTGAAAGATGACCCAACTTGC AAAAC 2003

QY 841 AAAGGAATTTTGGGCCACCAATTTTGGATTAGCAGGTCCTCTCTCTGACTATATCTCTCC 900
DB 2004 AAAGGAATTTTGGGCCACCAATTTTGGATTAGCAGGTCCTCTCTCTGACTATATCTCTCC 2063

QY 901 AGAAATTTTCTCTTTTGAAGTACTACTGGAATTTACATTTGATGGATGCTCTACAAGCC 960
DB 2064 AGAAATTTTCTCTTTTGAAGTACTACTGGAATTTACATTTGATGGATGCTCTACAAGCC 2123

QY 961 TCATGATCTACAGCCCTGGAAGAAATATCTACTGTGCTGTTCATATATATGGTGTGTC --- 1016
DB 2124 TCATGATCTACAGCCCTGGAAGAAATATCTACTGTGCTGTTCATATATATGGTGTGCGGT 2183

QY 1017 ----- 1016
DB 2184 CAAATAGAAATTTGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTTC 2243

QY 1017 ----- 1016
DB 2244 ATTGACTTAGATCGTGTGGGCATCCACGGCTGGTCTTATGAGGATACCTCTCCCTCATG 2303

QY 1017 -----CTCAGGTTGCTATTGCTGGGGCCCGAGTCACCTCTG 1051
DB 2304 GCATTAATGCAGAGGTCAGATATCTTCAGSGTTCCTATTGCTGGGGCCCGAGTCACCTCTG 2363

QY 1052 TCGATCTTCTATGATACAGGATACAGGAACTGTTATATGCGGTACCTGACCCAGCAAGTAA 1111
DB 2364 TCGATCTTCTATGATACAGGATACAGGAACTGTTATATGCGGTACCTGACCCAGCAAGTAA 2423

QY 1112 CAGGCTATTACTTAGGATCTGTGGCATGCAAGCAAGAAAGTTCCCTCTGACCAAT 1171
DB 2424 CAGGCTATTACTTAGGATCTGTGGCATGCAAGCAAGAAAGTTCCCTCTGACCAAT 2483

QY 1172 CGTTTACTGCTCTTACATGTTTCTGATGAGAAATGTCATTTTGCACATACCAATATA 1231
DB 2484 CGTTTACTGCTCTTACATGTTTCTGATGAGAAATGTCATTTTGCACATACCAATATA 2543

QY 1232 TTACTGAGTTTTTATGAGGCTCGAAAGCCATATGATTTACAGATCTATFCTCAGGAG 1291
DB 2544 TTACTGAGTTTTTATGAGGCTCGAAAGCCATATGATTTACAGATCTATFCTCAGGAG 2603

QY 1292 AGACACAGCATAGAGTTCCTGAATCGGGAAGAACATTTATGAACCTGCATCTTTTGCATAC 1351
DB 2604 AGACACAGCATAGAGTTCCTGAATCGGGAAGAACATTTATGAACCTGCATCTTTTGCATAC 2663

QY 1352 CTTCAAGAAACCTTTGGATCAGTATGCTGCTCTTAAAGTGATATAATTTTGAACCTGTG 1411
DB 2664 CTTCAAGAAACCTTTGGATCAGTATGCTGCTCTTAAAGTGATATAATTTTGAACCTGTG 2723

QY 1412 TAGAATCTCTGATATACACTGCTTATTTAAACCAATGAGGAGTTTAAATCAACAGAAAA 1471
DB 2724 TAGAATCTCTGATATACACTGCTTATTTAAACCAATGAGGAGTTTAAATCAACAGAAAA 2783

QY 1472 CACAGAAATGATCATCATCATTTTGTATCTGCTGATTAACATCTACTCTCTGAAATATAAT 1531
DB 2784 CACAGAAATGATCATCATCATTTTGTATCTGCTGATTAACATCTACTCTCTGAAATATAAT 2843

QY 1532 GTGGTCCATGAGGGCTCTACGGTTTGTGGTAGTAATCTAATACCTTAACCCACATGC 1591
DB 2844 GTGGTCCATGAGGGCTCTACGGTTTGTGGTAGTAATCTAATACCTTAACCCACATGC 2903

QY 1592 TCAAAATCAATATGATCATATTTCTGAGNAGCCAGCAATACCATAGAAATTAATTAATAA 1651
DB 2904 TCAAAATCAATATGATCATATTTCTGAGNAGCCAGCAATACCATAGAAATTAATTAATAA 2963

QY 1652 AAAAAAAAAAAAAA 1668
DB 2964 AAAAAAAAAAAAAA 2980

RESULT 7
ACA92421
ID ACA92421 standard; DNA; 2952 BP.
XX

AC ACA92421;
 XX 15-JUL-2003 (first entry)
 XX
 DE DNA encoding human PMMM-6.
 XX
 KW Human; protein modification and maintenance molecule; PMMM; cancer;
 KW cell proliferation disorder; atherosclerosis; neurological disorder;
 KW epilepsy; Huntington's disease; stroke; immune disorder; allergy;
 KW inflammatory disorder; AIDS; developmental disorder; hypothyroidism;
 KW Cushing's syndrome; gastrointestinal disorder; epithelial disorder;
 KW infection; cycostatic; antiarteriosclerotic; anticonvulsant; nootropic;
 KW neuroprotective; cerebroprotective; anti-HIV; antiallergic; vulnerary;
 KW antiinflammatory; thyromimetic; gene; ds.
 XX
 OS Homo sapiens.
 XX
 XX WO2003031939-A2.
 XX
 PD 17-APR-2003.
 XX
 XX 11-OCT-2002; 2002WO-US032850.
 XX
 XX 12-OCT-2001; 2001US-0329689P.
 PR 25-OCT-2001; 2001US-0335703P.
 PR 09-NOV-2001; 2001US-0348887P.
 PR 28-NOV-2001; 2001US-0334145P.
 PR 06-DEC-2001; 2001US-0337451P.
 PR 14-DEC-2001; 2001US-0340584P.
 XX
 XX (INCY-) INCYTE GENOMICS INC.
 XX
 XX Ramkumar J, Gorvad AE, Baughn MR, Emerling BM, Yang J, Lee SY;
 PI Tran UK, Bécha SD, Duggan BM, Lee EA, Griffin JA, Li JX;
 PI Sprague WW, Hafalia AJA, Chawla NK, Lehr-Mason PM, Kable AE, Yue H;
 PI Marquis JP, Yao MG, Richardson TW, Tang TY, Jin P, Chien D;
 PI Bhatia U, Burrill JJ, Lee S, Blake JJ, Ho A, Zheng W;
 XX
 XX WPI; 2003-430274/40.
 DR P-PSDB; ABU92026.
 XX
 XX New human protein modification and maintenance molecules (PMMM), useful
 PT for diagnosing, treating and preventing diseases or conditions associated
 PT with the aberrant PMMM expression e.g. cancer, atherosclerosis, or
 PT infections.
 XX
 XX Claim 5; Page 285-286; 31pp; English.
 XX
 XX The present invention relates to the isolation of human protein
 CC modification and maintenance molecules (PMMM), and the polynucleotide
 CC sequences encoding them. A total of 40 PMMM polypeptides (designated PMMM
 CC -1 to PMMM-40) are disclosed. The sequences of the invention are useful
 CC for diagnosing a condition or disease associated with the expression of
 CC PMMM in a subject, preparing a polyclonal or monoclonal antibody, and
 CC generating an expression profile of a sample containing the
 CC polynucleotides. The diseases or conditions associated with decreased
 CC expression or overexpression of PMMM are cell proliferation disorders
 CC (e.g. cancer, atherosclerosis), neurological disorders (e.g. epilepsy,
 CC Huntington's disease, stroke), immune/inflammatory disorders, (e.g. AIDS,
 CC allergies), developmental disorders (e.g. hypothyroidism, Cushing's
 CC syndrome), gastrointestinal or epithelial disorders, and infections. The
 CC PMMM polypeptides or their fragments are useful in screening compounds
 CC for effectiveness as agonists or antagonists of the polypeptides, or in
 CC altering the expression of the target polynucleotide and compounds that
 CC specifically bind to, or modulate the activity of the polypeptide.
 CC ACA92416-ACA92455 encode the human PMMM polypeptides of the invention
 XX
 XX Sequence 2952 BP; 886 A; 604 C; 664 G; 798 T; 0 U; 0 Other;
 SQ
 Query Match 89.5%; Score 1493.4; DB 10; Length 2952;
 Best Local Similarity 91.8%; Pred. No. 0;
 Matches 1651; Conservative 0; Mismatches 1; Indels 147; Gaps 1;

QY 1 AACAGGTACAGCAAAATCCCTAAAGTCACCTTTTAAGATGTCAGAAATATGATGCTGCTGA 60
 DB 1154 AACAGGTACAGCAAAATCCCTAAAGTCACCTTTTAAGATGTCAGAAATATGATGCTGCTGA 1213
 QY 61 AGGAAGGATCATAGATGTCATAGATAAGAACTAAATCAACCTTTTCAGATTCCTATTGGA 120
 DB 1214 AGGAAGGATCATAGATGTCATAGATAAGAACTAAATCAACCTTTTCAGATTCCTATTGGA 1273
 QY 121 AGGAGTTGAATATATTTGCCAGAGCTGGATGAGCTCCTGAGGGAAAAATATGCTGGTCCAT 180
 DB 1274 AGGAGTTGAATATATTTGCCAGAGCTGGATGAGCTCCTGAGGGAAAAATATGCTGGTCCAT 1333
 QY 181 CCTACTAGATCGCTCCAGACTCGCCTACAGATAGTGTGATCTCACCTGAATATTATTAT 240
 DB 1334 CCTACTAGATCGCTCCAGACTCGCCTACAGATAGTGTGATCTCACCTGAATATTATTAT 1393
 QY 241 CCCAGTAGAAGATGATGTTATGGAAGGCAGAGACTCAATTCAGTCAGTCGCTGATTCCTGT 300
 DB 1394 CCCAGTAGAAGATGATGTTATGGAAGGCAGAGACTCAATTCAGTCAGTCGCTGATTCCTGT 1453
 QY 301 GACGCCACTAATATCTATGAAAGAAACCAACAGACATCTGATAAATATCCATGACATCTT 360
 DB 1454 GACGCCACTAATATCTATGAAAGAAACCAACAGACATCTGATAAATATCCATGACATCTT 1513
 QY 361 TCATGTTTTTCCCAAGTCACGAAGAGAAATGAGTTATTTTGGCTCTGGAATGCAA 420
 DB 1514 TCATGTTTTTCCCAAGTCACGAAGAGAAATGAGTTATTTTGGCTCTGGAATGCAA 1573
 QY 421 AACAGGTTTTCCGTCATTTATACAAATTTACATCTATTTTAAAGGAAGCAATATAACG 480
 DB 1574 AACAGGTTTTCCGTCATTTATACAAATTTACATCTATTTTAAAGGAAGCAATATAACG 1633
 QY 481 ATCCAGTGTGGGCTGCCTGCCCTCCCAAGTCATTTCAAAGTGTCTCTATCAAGAGGAGATAGC 540
 DB 1634 ATCCAGTGTGGGCTGCCTGCCCTCCCAAGTCATTTCAAAGTGTCTCTATCAAGAGGAGATAGC 1693
 QY 541 AATTACAGTGTGTAATGGGAAGTTCTTGGCGGCGATGGATCTAATATCCAAAGTTGATGA 600
 DB 1694 AATTACAGTGTGTAATGGGAAGTTCTTGGCGGCGATGGATCTAATATCCAAAGTTGATGA 1753
 QY 601 AGTCAGAGGCTGGTATATTTTGAAGGNCACCAAGACTCCCTTTTAGAGCATCACCTGTA 660
 DB 1754 AGTCAGAGGCTGGTATATTTTGAAGGNCACCAAGACTCCCTTTTAGAGCATCACCTGTA 1813
 QY 661 CGTAGTCAGTTACGTAATCCCTGGAGAGGTGACAAAGGCTGACTGACCGTGGCTACTCACA 720
 DB 1814 CGTAGTCAGTTACGTAATCCCTGGAGAGGTGACAAAGGCTGACTGACCGTGGCTACTCACA 1873
 QY 721 TTCTTGTGTCATCAGTCAGCAGCTGTGACTTTCTTTATAAGTAAGTATAGTAACCAAGAGAA 780
 DB 1874 TTCTTGTGTCATCAGTCAGCAGCTGTGACTTTCTTTATAAGTAAGTATAGTAACCAAGAGAA 1933
 QY 781 TCCACACTGTGTGCTCCCTTTTCAAGCTATCAAGTCTCAAGATGACCCCACTTCGCAAAAC 840
 DB 1934 TCCACACTGTGTGCTCCCTTTTCAAGCTATCAAGTCTCAAGATGACCCCACTTCGCAAAAC 1993
 QY 841 AAAGGAATTTTGGGCCACCAATTTTGGATTTCAGCAGGTCCTCTTCTGACTATATCTCCTCC 900
 DB 1994 AAAGGAATTTTGGGCCACCAATTTTGGATTTCAGCAGGTCCTCTTCTGACTATATCTCCTCC 2053
 QY 901 AGAAATTTTCTCTTTTGAAGAGTACTACTGGAATTTACATTTGATGGATGCTCTCAAGGCC 960
 DB 2054 AGAAATTTTCTCTTTTGAAGAGTACTACTGGAATTTACATTTGATGGATGCTCTCAAGGCC 2113
 QY 961 TCATGATCTACAGCTCGGAAGAAATATCTTACTGCTGTGTTTCATATATATGATGCTGCTCA 1020
 DB 2114 TCATGATCTACAGCTCGGAAGAAATATCTTACTGCTGTGTTTCATATATATGATGCTGCTCA 2173
 QY 1021 ----- 1020
 DB 2174 GGTGCAAGTGGTGAATATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTTAGC 2233
 QY 1021 ----- 1020

Db 2234 CTCTCTAGTGTATGTGTTGTAGTATGATAGACAAACAGGGGATCTGTGTCACCGAGGCTTAA 2293
Qy 1021 -----GGTTGCTATTGCTGGGGCCCCAGTCACTCTGTG 1053
Db 2294 ATTGGAAGCGGCTTTAAATATAAATAGTTGCTATTGCTGGGGCCCCAGTCACTCTGTG 2353
Qy 1054 GATCTTCTATGATACAGGATACACGGAACGTTTATATGGGTACCCCTGACCCAGAAATGAACA 1113
Db 2354 GATCTTCTATGATACAGGATACACGGAACGTTTATATGGGTACCCCTGACCCAGAAATGAACA 2413
Qy 1114 GGGCTATTACTTAGGATCTGGGCCATGCAAGCAGAAAGGTTCCCTCTGAAACCAATCG 1173
Db 2414 GGGCTATTACTTAGGATCTGGGCCATGCAAGCAGAAAGGTTCCCTCTGAAACCAATCG 2473
Qy 1174 TTTACTGCTCTTACATGGTTTCTGGATGAGAATGTCATTTTGGCATACACCATATAT 1233
Db 2474 TTTACTGCTCTTACATGGTTTCTGGATGAGAATGTCATTTTGGCATACACCATATAT 2533
Qy 1234 ACTGAGTTTTTTAGTGAGGGCTGGAAAGCCATATGATTTACAGATCTATCCTCAGGAGAG 1293
Db 2534 ACTGAGTTTTTTAGTGAGGGCTGGAAAGCCATATGATTTACAGATCTATCCTCAGGAGAG 2593
Qy 1294 ACACAGCATAAAGGTTCTGGAATCGGAGAACATTATGAACCTGTCATCTTTTGCACCTACCT 1353
Db 2594 ACACAGCATAAAGGTTCTGGAATCGGAGAACATTATGAACCTGTCATCTTTTGCACCTACCT 2653
Qy 1354 TCAAGAAACCTTTGGATCAGCTATTGCTGCTCTAAAGTGATATAATTTTGACCTGTGA 1413
Db 2654 TCAAGAAACCTTTGGATCAGCTATTGCTGCTCTAAAGTGATATAATTTTGACCTGTGA 2713
Qy 1414 GAACCTCTGGTATACATGCTGCTATTTAAACCAATGAGGAGTTTAACTCAACAGAAACA 1473
Db 2714 GAACCTCTGGTATACATGCTGCTATTTAAACCAATGAGGAGTTTAACTCAACAGAAACA 2773
Qy 1474 CAGAAATGATCATCATTTTGTATGATACCTGCGCATGTAACATCTACTCCTGAAATAAATG 1533
Db 2774 CAGAAATGATCATCATTTTGTATGATACCTGCGCATGTAACATCTACTCCTGAAATAAATG 2833
Qy 1534 GTGCGCATGAGGGGTCTACGGTTTGTGGTAGTAATCTAATACCTTAAACCCACATGCTC 1593
Db 2834 GTGCGCATGAGGGGTCTACGGTTTGTGGTAGTAATCTAATACCTTAAACCCACATGCTC 2893
Qy 1594 AAAATCAAATGATACATATTCCTGAGAGACCCAGCAATACCATAAGAAATTAATAAAAA 1652
Db 2894 AAAATCAAATGATACATATTCCTGAGAGACCCAGCAATACCATAAGAAATTAATAAAAA 2952

RESULT 8
AAH15009
ID AAH15009 standard; cDNA; 2161 BP.
XX
AC AAH15009;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human cDNA sequence SEQ ID NO:12963.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-00116126.
XX
PR 29-JUL-1999; 99JP-00248036.
PR 27-AUG-1999; 99JP-00300253.
PR 11-JAN-2000; 2000JP-00118776.
PR 02-MAY-2000; 2000JP-00183767.
PR 09-JUN-2000; 2000JP-00241899.

XX (HELI-) HELIX RES INST.
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI; 2001-318749/34.
XX
XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.
XX
XX Claim 8; SEQ ID NO 12963; 2537pp + Sequence Listing; English.
XX
XX The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides; and the combination of the 5'-end sequence and 3'-end sequence is selected from those defined in the specification. The primer sets are useful for synthesizing polynucleotides, particularly full-length cDNAs. The primers are also useful for the gene therapy. The primers are useful for synthesizing polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention
XX
XX Sequence 2161 BP; 662 A; 429 C; 452 G; 618 T; 0 U; 0 Other;
Qy Query Match 89.1%; Score 1487.4; DB 4; Length 2161;
Best Local Similarity 91.7%; Pred. No. 0;
Matches 1645; Conservative 0; Mismatches 1; Indels 147; Gaps 1;
Qy 1 AACAGGTACAGCAAAATCCTAAAGTCACCTTTTAAGATGTCAGAAATATGATGCTGA 60
Db 369 AACAGGTACAGCAAAATCCTAAAGTCACCTTTTAAGATGTCAGAAATATGATGCTGA 428
Qy 61 AGGAAGGATCATAGATGTCATAGATAAGGAACCTTAATCAACCTTTTGAGATTCTATTGA 120
Db 429 AGGAAGGATCATAGATGTCATAGATAAGGAACCTTAATCAACCTTTTGAGATTCTATTGA 488
Qy 121 AGGAGTTGAATATATTGCCAGAGCTGGATGACCTCCTGAGGAAATATGCTGGTCCAT 180
Db 489 AGGAGTTGAATATATTGCCAGAGCTGGATGACCTCCTGAGGAAATATGCTGGTCCAT 548
Qy 181 CCTACTAGATCGCTCCAGACTCGCTACAGATGTTGATCTCACTGCAATATTAT 240
Db 549 CCTACTAGATCGCTCCAGACTCGCTACAGATGTTGATCTCACTGCAATATTAT 608
Qy 241 CCCAGTAGAAGATGATGTTATGGAAAGCGAGACTCAATTGAGTCTAGTCTGATCTGT 300
Db 609 CCCAGTAGAAGATGATGTTATGGAAAGCGAGACTCAATTGAGTCTAGTCTGATCTGT 668
Qy 301 GACGCCACTAATTTATCTATGAAGAAACAAACAGACATCTGGATAAATATCCATGACATCT 360
Db 669 GACGCCACTAATTTATCTATGAAGAAACAAACAGACATCTGGATAAATATCCATGACATCT 728
Qy 361 TCATGTTTTTCCCAAGTCAAGAGGAAATTTAGTTTATTTTTTGCCTCTGAATGCAA 420
Db 729 TCATGTTTTTCCCAAGTCAAGAGGAAATTTAGTTTATTTTTTGCCTCTGAATGCAA 788
Qy 421 AACAGGTTTCCGTCATTTATACAAAATTTACATCTATTTTAAAGGAAAGCAATATATAACG 480

Db 951 AACAGGTACAGCAAACTCTAAAGTCACCTTTTAAAGATGTCAGAAATATGATTGATGCTGA 1010
Qy 61 AGAAGGATCATAGATGTCATAGATAAGGAACCTTAATCAACCTTTTCAGATTCTATTGGA 120
Db 1011 AGAAGGATCATAGATGTCATAGATAAGGAACCTTAATCAACCTTTTCAGATTCTATTGGA 1070
Qy 121 AGAGATTGAATATATTGCGCAGAGCTGGATGGAGCTCCCTGAGGGAAATATGCTGGTCCAT 180
Db 1071 AGAGATTGAATATATTGCGCAGAGCTGGATGGAGCTCCCTGAGGGAAATATGCTGGTCCAT 1130
Qy 181 CCTACTAGATCCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAATATTATAT 240
Db 1131 CCTACTAGATCCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAATATTATAT 1190
Qy 241 CCAGTAGAAGATGATGTTATGGAAGGCGAGAGACTCAATTGAGTCAGTGCCTGATCTGT 300
Db 1191 CCAGTAGAAGATGATGTTATGGAAGGCGAGAGACTCAATTGAGTCAGTGCCTGATCTGT 1250
Qy 301 GAGCCACTAATATTCTATGAAAGAAACAACAGACATCTGGATAAATATCCATGACATCTT 360
Db 1251 GAGCCACTAATATTCTATGAAAGAAACAACAGACATCTGGATAAATATCCATGACATCTT 1310
Qy 361 TCATGTTTTTCCCAAGTCACGAAAGAGGAAATGAGTATTATTTTGGCTCTGAAATGCAA 420
Db 1311 TCATGTTTTTCCCAAGTCACGAAAGAGGAAATGAGTATTATTTTGGCTCTGAAATGCAA 1370
Qy 421 AACAGGTTTTCCGTCATTTATACAAATATACATCTATTATTAAGGAAGCAATATAACG 480
Db 1371 AACAGGTTTTCCGTCATTTATACAAATATACATCTATTATTAAGGAAGCAATATAACG 1430
Qy 481 ATCCAGTGTGGGCTGCCTGCCCTCCAAAGTGATTTTCAAGTGTCTTCAAGAGAGATAGC 540
Db 1431 ATCCAGTGTGGGCTGCCTGCCCTCCAAAGTGATTTTCAAGTGTCTTCAAGAGAGATAGC 1490
Qy 541 AATTACAGTGTGATGGAGTCTTTGGCCGCATGGATCTAATATCCAGTTGATCA 600
Db 1491 AATTACAGTGTGATGGAGTCTTTGGCCGCATGGATCTAATATCCAGTTGATCA 1550
Qy 601 AGTCAGAAGCTGTATATTGAGGCAACAAGACTCCCTTTAGAGCATCACTGTA 660
Db 1551 AGTCAGAAGCTGTATATTGAGGCAACAAGACTCCCTTTAGAGCATCACTGTA 1610
Qy 661 CGTAGTCAGTTAGTAATCTTGGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCA 720
Db 1611 CGTAGTCAGTTAGTAATCTTGGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCA 1670
Qy 721 TTCTTGCTGCATCAGTCAGCATGTGACTTCTTTTAAAGTAAGTATAGTAACAGAGAA 780
Db 1671 TTCTTGCTGCATCAGTCAGCATGTGACTTCTTTTAAAGTAAGTATAGTAACAGAGAA 1730
Qy 781 TCACACTGTGTGTCCTTTTAAAGCTATCAAGTATCAAGTATGAAAGATGACCCAACTTGCAAAAC 840
Db 1731 TCACACTGTGTGTCCTTTTAAAGCTATCAAGTATCAAGTATGAAAGATGACCCAACTTGCAAAAC 1790
Qy 841 AAAGGAATTTTGGGCCACCATTTTGGATTTCAGCAGGTCCTCTTCTCTGACTACTCTCC 900
Db 1791 AAAGGAATTTTGGGCCACCATTTTGGATTTCAGCAGGTCCTCTTCTCTGACTACTCTCC 1850
Qy 901 AGAAATTTTCTCTTTTGAAGTACTACTTGGATTTACATTTGATGGGATGCTCTACAGCC 960
Db 1851 AGAAATTTTCTCTTTTGAAGTACTACTTGGATTTACATTTGATGGGATGCTCTACAGCC 1910
Qy 961 TCATGATCTACAGCTGGAAGAAATATCTTACTGTGCTGCTCATATATGTTGCTTCA 1020
Db 1911 TCATGATCTACAGCTGGAAGAAATATCTTACTGTGCTGCTCATATATGTTGCTTCA 1970
Qy 1021 GGTTCCTATTGTGGGGCCCGCAGTCACCTCTGTGGATCTTCTATGATACAGGATACACGA 1080
Db 1971 GGTTCCTATTGTGGGGCCCGCAGTCACCTCTGTGGATCTTCTATGATACAGGATACACGA 2030
Qy 1081 AGTTTATAGGTCACCTTACACAGATGACAGGCTATTACTAGGATCTGTGGCAT 1140
Db 2031 AGTTTATAGGTCACCTTACACAGATGACAGGCTATTACTAGGATCTGTGGCAT 2090

Qy 1141 GCAAGCAGAAAGTTCCCTCTGAAACCAATCGTTTACTGCTTCTACATGTTTCTCTGGA 1200
Db 2091 GCAAGCAGAAAGTTCCCTCTGAAACCAATCGTTTACTGCTTCTACATGTTTCTCTGGA 2150
Qy 1201 TGAGATGTCATTTTGGACATACACAGTATATTACTGAGTATTTTGTAGTGGGCTGAAA 1260
Db 2151 TGAGATGTCATTTTGGACATACACAGTATATTACTGAGTATTTTGTAGTGGGCTGAAA 2210
Qy 1261 GCCATATGATTTACAGATCTATCTCTCAGAGAGACACAGCATAAAGTTCCTGAATCGGG 1320
Db 2211 GCCATATGATTTACAGATCTATCTCTCAGAGAGACACAGCATAAAGTTCCTGAATCGGG 2270
Qy 1321 AGAACATTTAGAACTGATCTTTTTCGACTACCTTCAAGAAAACTTGGATCACGTAATTC 1380
Db 2271 AGAACATTTAGAACTGATCTTTTTCGACTACCTTCAAGAAAACTTGGATCACGTAATTC 2330
Qy 1381 TGCTCTAAAAAGTGATATAA 1399
Db 2331 TGCTCTAAAAAGTGATATAA 2349

RESULT 10
AAH99934
ID AAH99934 standard; cdna; 3143 bp.
XX AAH99934;
XX
XX 12-APR-2002 (first entry)
XX
XX cDNA encoding 21953 human prollyl oligopeptidase.
XX
XX 21953 prollyl oligopeptidase; human; proline; endopeptidase; cancer;
XX cardiovascular disease; autoimmune disease; atopic allergy;
XX neuronal disorder; vascular disorder; prostate disorder; cystostatic;
XX antidiabetic; antichratic; antiasthmatic; antiinflammatory;
XX diabetes mellitus; arthritis; multiple sclerosis; asthma;
XX Grave's disease; neuronal disorder; demyelinating disease; ss.
XX Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 229..2877
FT /*tag= a
FT /product= "21953 prollyl oligopeptidase"
FT /note= "This region is specifically claimed in claim 2"
XX
XX WO200179473-A2.
XX
XX 25-OCT-2001.
XX
XX 11-APR-2001; 2001WO-US040483.
XX
XX 18-APR-2000; 2000US-0197508P.
XX
XX (MILL-) MILLENNIUM PHARM INC.
XX
XX Meyers RA, Williamson M;
XX
XX WPI; 2002-034353/04.
XX P-PSDB; AAG78415.
XX
XX New polypeptides 21953, member of human prollyl oligopeptidase family,
XX useful as diagnostic targets and therapeutic agents for controlling
XX cancer, lymphoma and leukemia.
XX
XX Claim 7; Page 100-102; 121pp; English.
XX
XX This invention relates to an isolated 21953 human prollyl oligopeptidase.
XX which is cystostatic, antidiabetic, antichratic, neuroprotective,
XX antichryoid, dermatological, antipsoriatic, antiasthmatic,
XX ophthalmological, antiinflammatory, neurotropic, antiparkinsonian,
XX anticonvulsant, gynaecological, vasodilating, cardiant,

CC antiatherosclerotic, anorectic and metabolic in its action. Uses include
CC gene therapy, expression or activity of 21953 protein modulator, it is
CC useful for identifying a compound which binds to it and can be used in
CC preventing, treating or detecting a cellular proliferative or
CC differentiative disorder. The 21953 molecules can act as novel diagnostic
CC targets and therapeutic agents for controlling disorders associated with
CC the aberrant activity or degradation of peptide hormones e.g., disorders
CC associated with cell differentiation and proliferation such as cancer,
CC immune function, reproductive, neurological and cardiovascular function.
CC The 21953 molecules are thus useful for treating and preventing cellular
CC proliferative and differentiative disorders, haematopoietic neoplastic
CC disorders, immune disorders such as autoimmune diseases, diabetes
CC mellitus, arthritis, multiple sclerosis, asthma, Grave's disease,
CC neuronal disorders, demyelinating diseases, vascular disorders and
CC metabolism or pain disorders. This sequence represents the cDNA encoding
CC sequence of 21953 human prolyl oligopeptidase
XX
SQ Sequence 3143 BP; 943 A; 644 C; 712 G; 844 T; 0 U; 0 Other;
Query Match 81.1%; Score 1353.4; DB 6; Length 3143;
Best Local Similarity 84.7%; Pred. No. 0;
Matches 1664; Conservative 0; Mismatches 1; Indels 300; Gaps 1;
QY 1 AACAGGTACAGCAAAATCCCTAAAGTCACCTTTTAAGATGCTCAGAAATTAATGATTGATGCTGA 60
DB 1179 AACAGGTACAGCAAAATCCCTAAAGTCACCTTTTAAGATGCTCAGAAATTAATGATTGATGCTGA 1238
QY 61 AGGAGGATCATAGATGTCATAGATNAGGAACCTAAATCAACCTTTTGAGATCTATTGGA 120
DB 1239 AGGAGGATCATAGATGTCATAGATNAGGAACCTAAATCAACCTTTTGAGATCTATTGGA 1298
QY 121 AGGAGTGAATATATATGCGAGAGCTGGATGACTCCCTGAGGGGAAAATATGCTTGGTCCAT 180
DB 1299 AGGAGTGAATATATATGCGAGAGCTGGATGACTCCCTGAGGGGAAAATATGCTTGGTCCAT 1358
QY 181 CCTACTAGATCGCTCCAGAGTCGCTACAGATAGTGTGATCTCACCTGGAATATTAT 240
DB 1359 CCTACTAGATCGCTCCAGAGTCGCTCGAGATAGTGTGATCTCACCTGGAATATTAT 1418
QY 241 CCCAGTAGAAGATGATGTTATGGAAGCGACAGACTCATTTGAGTCAGTGCCTGATTCGT 300
DB 1419 CCCAGTAGAAGATGATGTTATGGAAGCGACAGACTCATTTGAGTCAGTGCCTGATTCGT 1478
QY 301 GAGGCCACTAAATTTCTATGAAGAAACAAACAGACATCTGGATAAATATCCATGACATCTT 360
DB 1479 GAGGCCACTAAATTTCTATGAAGAAACAAACAGACATCTGGATAAATATCCATGACATCTT 1538
QY 361 TCATGTTTTTCCCAAGTCACGAAGAGGAAATTTGAGTTATTTTTTGGCTCTCGAATGCAA 420
DB 1539 TCATGTTTTTCCCAAGTCACGAAGAGGAAATTTGAGTTATTTTTTGGCTCTCGAATGCAA 1598
QY 421 AACAGGTTTCCGTCATTATACAAATTTACATCTATTTTAAAGGAAGCAAAATATAACG 480
DB 1599 AACAGGTTTCCGTCATTATACAAATTTACATCTATTTTAAAGGAAGCAAAATATAACG 1658
QY 481 ATCCAGTGTGGGCTGCCCTGCTCCCAAGTGATTTCAAGTGTCTCTATCAAGAGGAGATAGC 540
DB 1659 ATCCAGTGTGGGCTGCCCTGCTCCCAAGTGATTTCAAGTGTCTCTATCAAGAGGAGATAGC 1718
QY 541 AATTAACAGTGTGGAATGGGAAGTTCTTGGCGGCGATGGATCTAATATCAAGTTGATGA 600
DB 1719 AATTAACAGTGTGGAATGGGAAGTTCTTGGCGGCGATGGATCTAATATCAAGTTGATGA 1778
QY 601 AGTCAGAGGCTGGTATATTTTGAAGGCACCAAGAGCTCCCTTTTACAGCATCACCTGTA 660
DB 1779 AGTCAGAGGCTGGTATATTTTGAAGGCACCAAGAGCTCCCTTTTACAGCATCACCTGTA 1838
QY 661 CGTAGTCAGTTACGTAAATCCCTGGAGAGGTGACAGGCTGACTGACCGTGGCTACTACA 720
DB 1839 CGTAGTCAGTTACGTAAATCCCTGGAGAGGTGACAGGCTGACTGACCGTGGCTACTACA 1898
QY 721 TTCTTGCTGCATCAGTCAGCATCTGTGACTTTTAAAGTAAGTATATAGTAAACCAAGAGAA 780

DB 1899 TTCTTGCTGCATCAGTCAGCATCTGTGACTTCTTTTATAAGTAAGTATAGTAAACCAAGAGAA 1958
QY 781 TCCACACTGTGTGTCCTCTTTTACAAGCTATCAAGTCTCTGAAGATGACCCCAACTTGCAGAAC 840
DB 1959 TCCACACTGTGTGTCCTCTTTTACAAGCTATCAAGTCTCTGAAGATGACCCCAACTTGCAGAAC 2018
QY 841 AAAGGAAATTTGGGCGACCATTTTGGATTGAGTTCAGAGGTCCTCTCTCTGACTATACCTCTCC 900
DB 2019 AAAGGAAATTTGGGCGACCATTTTGGATTGAGTTCAGAGGTCCTCTCTCTGACTATACCTCTCC 2078
QY 901 AGAAATTTCTCTTTTGAAGTACTACTGGATTTCATCTGATTTTACATTTGATGGATGCTCTTACAAGCC 960
DB 2079 AGAAATTTCTCTTTTGAAGTACTACTGGATTTCATCTGATTTTACATTTGATGGATGCTCTTACAAGCC 2138
QY 961 TCATGATCTTACAGCCTCGAAAGAAATATCTCTACTGCTCTCTCATATATATGATGGTGGTCTCA 1020
DB 2139 TCATGATCTTACAGCCTCGAAAGAAATATCTCTACTGCTCTCTCATATATATGATGGTGGTCTCA 2198
QY 1021 ----- 1020
DB 2199 GGTGCAGTTGGTGAATAATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCCCTAGC 2258
QY 1021 ----- 1020
DB 2259 CTCTCTAGTTATGTGTTGTAGTATAGACAAACAGGGGATCCTGTCAACCGAGGGCTTAA 2318
QY 1021 ----- 1020
DB 2319 ATTTGAAGCGCCTTTAAATATATAATATGAGTCAAAATAGAAAATTGACGATCAGGTGGAGG 2378
QY 1021 ----- 1020
DB 2379 ACTCCAATATCTAGCTTCTCGATATGATTTCAITGACTTATGATCGTGGGCGATCCACGG 2438
QY 1021 ----- 1020
DB 2439 CTGGTCTATGAGGATACCTCTCCCTGATGGCATTAATGCAGAGGTGAGATATCTTCAG 2498
QY 1021 GGTGTCTATTTGCTGGGGCCCCAGTCACTCTGTGATCTTTTATGATACAGGATACACGGA 1080
DB 2499 GGTGTCTATTTGCTGGGGCCCCAGTCACTCTGTGATCTTTTATGATACAGGATACACGGA 2558
QY 1081 ACCTTATATGGTTCACCTGACCAAGTAAACAGGGCTATTACTTAGATCTGTGGCCAT 1140
DB 2559 ACCTTATATGGTTCACCTGACCAAGTAAACAGGGCTATTACTTAGATCTGTGGCCAT 2618
QY 1141 GCAAGCAGAAAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGTTTCTCTGGA 1200
DB 2619 GCAAGCAGAAAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGTTTCTCTGGA 2678
QY 1201 TGGAATGTCCATTTTGGCATAATCCAGTATATTAATGATGATTTTGTAGGGGCTGGAAA 1260
DB 2679 TGGAATGTCCATTTTGGCATAATCCAGTATATTAATGATGATTTTGTAGGGGCTGGAAA 2738
QY 1261 GCATATGATTTTACAGATCTATCTCCAGAGAGACACAGATTAAGAGTTCTCTGAATCGGG 1320
DB 2739 GCATATGATTTTACAGATCTATCTCCAGAGAGACACAGATTAAGAGTTCTCTGAATCGGG 2798
QY 1321 AGAACATTTATGAACCTCATCTTTTGCATCACTCCCTCAAGAAAAACCTTGGATCACGATATGC 1380
DB 2799 AGAACATTTATGAACCTCATCTTTTGCATCACTCCCTCAAGAAAAACCTTGGATCACGATATGC 2858
QY 1381 TGCTCTAAAAGTGATATAAATTTTGGACCTGTGTGAGAACTCTCTGGTATACACTGGCTATTT 1440
DB 2859 TGCTCTAAAAGTGATATAAATTTTGGACCTGTGTGAGAACTCTCTGGTATACACTGGCTATTT 2918
QY 1441 AACCAATGAGGAGGTTTAAATCAACAGAAAAACAGAAATTCATCATCATTTTGTATACC 1500
DB 2919 AACCAATGAGGAGGTTTAAATCAACAGAAAAACAGAAATTCATCATCATTTTGTATACC 2978
QY 1501 TGCCATGTAACTACTCTCTGAAAAATAAATGTTGGTGGCCATGCAGGGGCTCTACGGTTGT 1560
DB 2979 TGCCATGTAACTACTCTCTGAAAAATAAATGTTGGTGGCCATGCAGGGGCTCTACGGTTGT 3038

Db 830 GGCCTCTTAAATATATAAATGGTCAAATAGAAATTTGACGATCAGGTGGAGGACTCAA 771
Qy 1021 ----- 1020
Db 770 TAATCTAGTCTTCGATATGATTTCAATGACTTGTAGATCGTGGGCATCCACGGCTGGTCC 711
Qy 1021 -----GGTTGCT 1027
Db 710 TATGGAGGATACCTCTCCCTGATGCGATTAATGACAGGTCAGATATCTTCAAGGTTGCT 651
Qy 1028 ATTGCTGGGCCCCAGTCACTCTGTGGATCTTATATGATACAGGATACAGGAACGTTAT 1087
Db 650 ATTGCTGGGCCCCAGTCACTCTGTGGATCTTATGATACAGGATACAGGAACGTTAT 591
Qy 1088 ATGGGTCACTCTGACAGGAATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCA 1147
Db 590 ATGGGTCACTCTGACAGGAATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCA 531
Qy 1148 GAAAGTTCCCTCTGAAACCAATGTTTACTGCTCTTACATGTTTCTCTGGATGGAAT 1207
Db 530 GAAAGTTCCCTCTGAAACCAATGTTTACTGCTCTTACATGTTTCTCTGGATGGAAT 471
Qy 1208 GTCCATTTTGACATACACAGTATATTACTAGTGTGAGGCTGGAAAGCCATAT 1267
Db 470 GTCCATTTTGACATACACAGTATATTACTAGTGTGAGGCTGGAAAGCCATAT 411
Qy 1268 GATTACAGATCTATCTCTCAGGAGAGACACAGCATAAGAGTTCCTGAATCGGAGAACAT 1327
Db 410 GATTACAGATCTATCTCTCAGGAGAGACACAGCATAAGAGTTCCTGAATCGGAGAACAT 351
Qy 1328 TATGAAGTCACTTTTGGTCACTACCTTCAAGAAACCTTGGATCAGCTATTGCTGCTTA 1387
Db 350 TATGAAGTCACTTTTGGTCACTACCTTCAAGAAACCTTGGATCAGCTATTGCTGCTTA 291
Qy 1388 AAGTGTATATAATTTTGACCTGTGTAGACTCTCTGGTATACACTGGCTATTTAAACCAA 1447
Db 290 AAGTGTATATAATTTTGACCTGTGTAGACTCTCTGGTATACACTGGCTATTTAAACCAA 231
Qy 1448 TGAGGAGGTTTAAATCAACAGAAACACACAGAAATTTGATCATCATCTTTGATACCTGGCCATG 1507
Db 230 TGAGGAGGTTTAAATCAACAGAAACACACAGAAATTTGATCATCATCTTTGATACCTGGCCATG 171
Qy 1508 TAAATCTACTCTCGAAATAAATGTGTGGCCATGACAGGGTCTACGGTTTGTGGTAGTA 1567
Db 170 TAAATCTACTCTCGAAATAAATGTGTGGCCATGACAGGGTCTACGGTTTGTGGTAGTA 111
Qy 1568 ATCTAATACCTTAACCCCATGCTCAAAATCAAAATGATATATATCTCTGAGAGCCAG 1627
Db 110 ATCTAATACCTTAACCCCATGCTCAAAATCAAAATGATATATATCTCTGAGAGCCAG 51
Qy 1628 CAATACATAAGAAATTTACTAAAAAATAAAAAAAAAAAAAAAAAAAAAA 1667
Db 50 CAATACATAAGAAATTTACTAAAAAATAAAAAAAAAAAAAAAAAAAAAA 11

RESULT 12
ABK83327
ID ABK83327 standard; cDNA; 4829 BP.
XX AC ABK83327;
XX DT 12-AUG-2002 (first entry)
XX DE cDNA encoding human DPRP-1 splice variant #3.
KW Human; serine protease; dipeptidyl peptidase IV-related protein; DPRP;
KW DPPIV; infection; human immunodeficiency virus; HIV-1; HIV-2; pain;
KW diabetes; infertility; obesity; anorexia; Parkinson's disease; stroke;
KW heart failure; hypertension; urinary retention; osteoporosis; cancer;
KW ulcer; allergy; cancer; psychotic disorder; neurological disorder;
KW dyskinesia; reproductive disorder; inflammatory disorder;
KW metabolic disorder; gene; ss.

OS Homo sapiens.
PN WO200231134-A2.
XX PD 18-APR-2002.
XX PF 12-OCT-2001; 2001WO-US031874.
XX PR 12-OCT-2000; 2000US-0240117P.
XX (FERR) FERRING BV.
PI Qi S, Akinsanya KO, Riviere PJ, Junien J;
XX WPI; 2002-444178/47.
DR P-PSDB; ABG61596.
XX New dipeptidyl peptidase IV-related proteins and nucleic acids encoding
PT the proteins, useful for treating e.g. fungal, bacterial, protozoan and
PT viral infections, cancers, allergies, neurological disorders, or pain.
XX Disclosure; Page 65-66; 113pp; English.
XX The present invention relates to the isolation of novel human serine
CC proteases referred to as dipeptidyl peptidase IV (DPPIV)-related proteins
CC (DPRP). The dipeptidyl peptidase IV-related proteins (DPRP) and nucleic
CC acids encoding them are useful for treating infections such as fungal,
CC bacterial, protozoan and viral infections, particularly infections caused
CC by human immunodeficiency virus (HIV-1 or HIV-2), pain, diabetes,
CC precocious puberty, infertility, obesity, anorexia, bulimia, Parkinson's
CC disease, acute heart failure, hypotension, hypertension, urinary
CC retention, osteoporosis, angina pectoris, stroke, ulcers, asthma,
CC allergies, cancers, migraine, vomiting, psychotic and neurological
CC disorders (e.g. anxiety, dementia, or schizophrenia), and dyskinesias.
CC These may also be used in discovering therapeutic agents for the
CC treatment of reproductive, inflammatory and metabolic disorders. ABK83322
CC -ABK83343 encode human DPRP proteins
XX SQ Sequence 4829 BP; 1466 A; 886 C; 1017 G; 1460 T; 0 U; 0 Other;
Query Match 80.9%; Score 1349.6; DB 6; Length 4829;
Best Local Similarity 84.5%; Pred. No. 0;
Matches 1664; Conservative 0; Mismatches 4; Indels 302; Gaps 1;
Qy 1 AACAGGTACAGCAAAATCCTAAAGTCACTTTTAAAGATGTGAGAAATTAATGATGCTGA 60
Db 1164 AACAGGTACAGCAAAATCCTAAAGTCACTTTTAAAGATGTGAGAAATTAATGATGCTGA 1223
Qy 61 AGGAAGATCATAGATGTGATAGATAGAACTAAATTCACCTTTTGAGATTCATTTGA 120
Db 1224 AGGAAGATCATAGATGTGATAGATAGAACTAAATTCACCTTTTGAGATTCATTTGA 1283
Qy 121 AGGATTTGATATATTTGCCAGAGCTGGATGCTCTCTGAGGAAATATGCTGGTCCAT 180
Db 1284 AGGATTTGATATATTTGCCAGAGCTGGATGCTCTCTGAGGAAATATGCTGGTCCAT 1343
Qy 181 CCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAATTTAT 240
Db 1344 CCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAATTTAT 1403
Qy 241 CCCAGTAGAAGATGATGTTTATGGAAGGACAGAGATCAATGAGTCAGTCGCTGATTCGT 300
Db 1404 CCCAGTAGAAGATGATGTTTATGGAAGGACAGAGATCAATGAGTCAGTCGCTGATTCGT 1463
Qy 301 GAGCCACTAATTTATCTATGAGAAACACAGACATCTGGATTAATATCCATCATCTT 360
Db 1464 GAGCCACTAATTTATCTATGAGAAACACAGACATCTGGATTAATATCCATCATCTT 1523
Qy 361 TCATGTTTTTCCCAAGTCAAGAGGAAATTCAGTTATTTTGTGCTCTGAATGCAA 420
Db 1524 TCATGTTTTTCCCAAGTCAAGAGGAAATTCAGTTATTTTGTGCTCTGAATGCAA 1593
Qy 421 AACAGGTTTCCGTCATTTTATACAAATTTACATCTATTTTAAAGGAAACAAATATAACG 480

1584	Db	 AACAGGTTCCGTCATTTATACAAATTTACATCTATTTTAAAGGAAGCAAAATAAAACG	1643
481	QY	ATPCAGTGGTGGCTGCTCTCCAGTGATTTCAAGTGTCTCTATCAAAGAGGAGATAGC	540
1644	Db	ATCCAGTGGTGGCTGCTCTCCAGTGATTTCAAGTGTCTCTATCAAAGAGGAGATAGC	1703
541	QY	AATTACCAGTGGTGAATGGGAAGTTCTTGGCCGGCATGGATCTAATATCCAAAGTTGATGA	600
1704	Db	AATTACCAGTGGTGAATGGGAAGTTCTTGGCCGGCATGGATCTAATATCCAAAGTTGATGA	1763
601	QY	AGTCAGAAGGCTGGTATATTTTGAAGSCACCAAGACTCCCTCTTTAGAGCATCACCTGTGA	660
1764	Db	AGTCAGAAGGCTGGTATATTTTGAAGSCACCAAGACTCCCTCTTTAGAGCATCACCTGTGA	1823
661	QY	CGTAGCTCAGTTACGTAAATCCTGGAGAGGTGACAGGCTGACTGCAGCTGGCTACTCACA	720
1824	Db	CGTAGCTCAGTTACGTAAATCCTGGAGAGGTGACAGGCTGACTGCAGCTGGCTACTCACA	1883
721	QY	TTCTTGTCTGCATCAGTCAGACACTGTCACCTCTTTATAGTAGTATAGTATACAGAGAA	780
1884	Db	TTCTTGTCTGCATCAGTCAGACACTGTCACCTCTTTATAGTAGTATAGTATACAGAGAA	1943
781	QY	TCCACACTGTGTGCCCTTTTCAAGCTATCAAGTCTCGAAGTGCACCCACTTGCACAAAC	840
1944	Db	TCCACACTGTGTGCCCTTTTCAAGCTATCAAGTCTCGAAGTGCACCCACTTGCACAAAC	2003
841	QY	AAAGGAATTTTGGGCCACCAATTTTGGATTACAGAGGCTCTCTTCTGCACTATACTCTCC	900
2004	Db	AAAGGAATTTTGGGCCACCAATTTTGGATTACAGAGGCTCTCTTCTGCACTATACTCTCC	2063
901	QY	AGAAATTTTCTCTTTTGAAGTACTACTCGATTTACATTGATGGGATGCTCTACAAGCC	960
2064	Db	AGAAATTTTCTCTTTTGAAGTACTACTCGATTTACATTGATGGGATGCTCTACAAGCC	2123
961	QY	TCATCATCTACAGCCTGGAAAGAAATATCTACTGTGCTGTTTATATATGGTGGTCTCA	1020
2124	Db	TCATCATCTACAGCCTGGAAAGAAATATCTACTGTGCTGTTTATATATGGTGGTCTCA	2183
1021	QY	-----	1020
2184	Db	CAGGTGCAGTTGGTGAATAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCCCTA	2243
1021	QY	-----	1020
2244	Db	GCCTCTCTAGGTTATGTGTTGTAGTGATAGACACAGGGGATCCTGTCAACGAGGCTT	2303
1021	QY	-----	1020
2304	Db	AAATTTGAAGCGCCTTTAAATATAAATGGGTCAATAGAAATTGACGATCAGGTGGAA	2363
1021	QY	-----	1020
2364	Db	GGACTCCAAATCTAGCTTCTCGATATGATTTCAATTGACTGTAGATCGTGGGCATCCAC	2423
1021	QY	-----	1020
2424	Db	GGCTGGTCCATGAGGATACCTCTCCCTGATGGCATTAATGACAGAGTCAGATATCTTC	2483
1021	QY	--GGTTGCTATTGCTGGGGCCCAAGTCACTCTGTGGATCTTCTATGATACAGGATACAG	1078
2484	Db	AGGGTTGCTATTGCTGGGGCCCAAGTCACTCTGTGGATCTTCTATGATACAGGATACAG	2543
1079	QY	GAACTGTTATGGGTCACCTGCACCAAGATGAACAGGCGCTATTACTTAGGATCTGTGGCC	1138
2544	Db	GAACTGTTATGGGTCACCTGCACCAAGATGAACAGGCGCTATTACTTAGGATCTGTGGCC	2603
1139	QY	ATGCAAGCAGAAAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGTTTCTCG	1198
2604	Db	ATGCAAGCAGAAAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGTTTCTCG	2663
1199	QY	GATGAGAAATGTCATTTTGACATACCAAGTATTAATCTAGTTTTTTTAGTGAGGCGTGA	1258

RESULT 13	
AAC85694	
ID	AAC85694 standard; cDNA; 3120 BP.
XX	
XX	AAC85694;
XX	
DT	29-JUN-2001 (first entry)
XX	
DE	Nucleotide sequence of human DPP8.
XX	
XX	Human; dipeptidyl aminopeptidase; DPP8; prolly oligopeptidase;
XX	dipeptidyl peptidase; DPPIV; T cell; cleavage; diarrhoea;
KW	growth hormone deficiency; glucose level; mucosal regeneration;
KW	non-insulin dependent diabetes mellitus; glucose intolerance;
KW	immunosuppression; ss.
XX	
OS	Homo sapiens.
XX	
PH	Key
XX	Location/Qualifiers
CDS	214..2862
FT	/*tag= a
FT	/product= "Human DPP8"
XX	
FN	WC200119866-A1.
XX	
PD	22-MAR-2001.
XX	
PF	11-SEP-2000; 2000WO-AU001085.
XX	
PR	10-SEP-1999; 99AU-00002762.
PR	18-FEB-2000; 2000AU-00005709.
XX	
PA	(UNSY) UNIV SYDNEY.
XX	
PI	Abbott CA, Gorell MD;
XX	
DR	WPI; 2001-281520/29.
DR	P-PSDB; AAB47187.
XX	
XX	New human dipeptidyl aminopeptidase (DPP8) useful for cleaving
PT	substrates, identifying inhibitors of DPP8 catalytic activity wh
PT	therapeutic uses, and for detecting activated T cells.


```
Db 2964 TGCCATGTAACATCTACTCCTGAAATAAATGCTGGTGCATGCGAGGGTCTACGGTTTGT 3023
Qy 1561 GGTAGTAATCTAATACCTTAAACCCACATGCTCAAAATCAATGATACATATTCCTGAGA 1620
Db 3024 GGTAGTAATCTAATACCTTAAACCCACATGCTCAAAATCAATGATACATATTCCTGAGA 3083
Qy 1621 GACCCAGCAATACCATAGAAATTAATAAAAAA 1657
Db 3084 GACCCAGCAATACCATAGAAATTAATAAAAAA 3120

RESULT 14
AAD38956
ID AAD38956 standard; cDNA; 3120 BP.
XX
AC AAD38956;
XX
DT 23-SEP-2002 (first entry)
XX
DE Human dipeptidyl peptidase 8 (DPP8) cDNA.
XX
KW Human; dipeptidyl peptidase; DPP; neoplasia; type II diabetes; cirrhosis;
KW autoimmunity; human immuno deficiency virus; HIV infection; cytostatic;
KW graft rejection; antidiabetic; antiinflammatory; immunosuppressive;
KW antiviral; enzyme; gene; ss.
XX
OS Homo sapiens.
XX
PH Key
FT CDS
FT /tag= a
FT /product= "Human DPP8 protein"
XX
PN WO200234900-A1.
XX
PD 02-MAY-2002.
XX
PF 29-OCT-2001; 2001WO-AU001388.
XX
PR 27-OCT-2000; 2000AU-00001078.
XX
PA (UNSY ) UNIV SYDNEY.
XX
PI Abbott CA, Gorrell MD;
XX
DR WPI; 2002-454646/48.
XX
DR P-PSDB; AAE24170.
XX
PT New dipeptidyl peptidase (DPP) peptides, useful for screening inhibitors
PT of DPP catalytic activity, which may be employed to treat e.g. neoplasia,
PT type II diabetes, cirrhosis, autoimmunity, graft rejection and HIV
PT infection.
XX
PS Example; Fig 1; 91pp; English.
XX
CC The present invention relates to dipeptidyl peptidase (DPP) proteins and
CC polynucleotides encoding such proteins. The DPP peptides are useful for
CC screening inhibitors of DPP catalytic activity. The inhibitors are useful
CC for treating neoplasia, type II diabetes, cirrhosis, autoimmunity, graft
CC rejection and HIV (human immuno deficiency virus) infection. The present
CC sequence is human DPP8 cDNA
XX
SQ Sequence 3120 BP; 936 A; 637 C; 706 G; 841 T; 0 U; 0 Other;

Query Match 80.7%; Score 1347; DB 6; Length 3120;
Best Local Similarity 84.7%; Pred. No. 0;
Matches 1657; Conservative 0; Mismatches 0; Indels 300; Gaps 1;

Qy 1 AACAGGTACAGCAATCTCAAGTCACCTTTTAAGATGTCAGAAATTAATGATGCTGCA 60
Db 1164 AACAGGTACAGCAATCTCAAGTCACCTTTTAAGATGTCAGAAATTAATGATGCTGCA 1223
Qy 61 AGGAAGGATCATAGATGTGATAGATAAGGAACCTAATTCACCTTTTGAGATTCATTGGA 120
```

```
Db 1224 AGGAAGGATCATAGATGTGATAGATAAGGAACCTAATTCACCTTTTGAGATTCATTGGA 1283
Qy 121 AGGAGTTGAATATATATTCGAGAGCTGGATGGACTCCTGAGGAAATATGCTTGGTCCAT 180
Db 1284 AGGAGTTGAATATATATTCGAGAGCTGGATGGACTCCTGAGGAAATATGCTTGGTCCAT 1343
Qy 181 CCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGCAATATTAT 240
Db 1344 CCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGCAATATTAT 1403
Qy 241 CCCAGTAGAAGATGATGTTATGAAAAGGCAGAGACTCAATTCAGTCAGTCGCTGATCTGT 300
Db 1404 CCCAGTAGAAGATGATGTTATGAAAAGGCAGAGACTCAATTCAGTCAGTCGCTGATCTGT 1463
Qy 301 GAGCCCACTAATTTATCTATGAAGAACAACAGACATCTGGATAAATATCCATGACATCTT 360
Db 1464 GAGCCCACTAATTTATCTATGAAGAACAACAGACATCTGGATAAATATCCATGACATCTT 1523
Qy 361 TCATGTTTTTCCCAAGTCACGAAGAGGAATTTAGTTTATTTTTCGCTCTGAATGCA 420
Db 1524 TCATGTTTTTCCCAAGTCACGAAGAGGAATTTAGTTTATTTTTCGCTCTGAATGCA 1583
Qy 421 AACAGGTTCCGTCATTTATACAAAATTTACATCTATTTTAAAGGAAGCAATATAACG 480
Db 1584 AACAGGTTCCGTCATTTATACAAAATTTACATCTATTTTAAAGGAAGCAATATAACG 1643
Qy 481 ATCCAGTGTGGCTGCTGCTCCAAAGTATTTCAAGTGTCTCTATCAAAAGAGAGATAGC 540
Db 1644 ATCCAGTGTGGCTGCTGCTCCAAAGTATTTCAAGTGTCTCTATCAAAAGAGAGATAGC 1703
Qy 541 AATTACAGTGTGGAATGGAAAGTTCTTTGGCCGCGATGATCTAATATCCAAGTTGATGA 600
Db 1704 AATTACAGTGTGGAATGGAAAGTTCTTTGGCCGCGATGATCTAATATCCAAGTTGATGA 1763
Qy 601 AGTCAGAAGCTGGTATATTTTGAAGGCACCAAGAGACTCCCTTTTAGAGCATCACCTGTA 660
Db 1764 AGTCAGAAGCTGGTATATTTTGAAGGCACCAAGAGACTCCCTTTTAGAGCATCACCTGTA 1823
Qy 661 CGTAGTCAGTTACGTAATCTCTGGAGAGGTCAGCAAGGCTGACTACCGTGGCTACTCACA 720
Db 1824 CGTAGTCAGTTACGTAATCTCTGGAGAGGTCAGCAAGGCTGACTACCGTGGCTACTCACA 1883
Qy 721 TTCTTGCTGCATCAGTCAGCACTGTGACTTCTTTAAGTAAGTATAGTAACCAAGAGAA 780
Db 1884 TTCTTGCTGCATCAGTCAGCACTGTGACTTCTTTAAGTAAGTATAGTAACCAAGAGAA 1943
Qy 781 TCCACACTGTGTGCTCCCTTTCAAGCTATCAAGTCTCGAAGATGACCCAACTTGCAAAAC 840
Db 1944 TCCACACTGTGTGCTCCCTTTCAAGCTATCAAGTCTCGAAGATGACCCAACTTGCAAAAC 2003
Qy 841 AAGGGAATTTTGGGCCACATTTTGGATTTCAGAGGTCCTCTTCTGACTATCTCTCC 900
Db 2004 AAGGGAATTTTGGGCCACATTTTGGATTTCAGAGGTCCTCTTCTGACTATCTCTCC 2063
Qy 901 AGAAATTTCTCTTTTGAAGTACTACTGGATTTTACATTTGTATGGGATGCTCTACAAGCC 960
Db 2064 AGAAATTTCTCTTTTGAAGTACTACTGGATTTTACATTTGTATGGGATGCTCTACAAGCC 2123
Qy 961 TCAATGATCTACAGCTGGAAAGAAATATCTTACTGTGCTGTTTCATATATATGGTGGTCTCA 1020
Db 2124 TCAATGATCTACAGCTGGAAAGAAATATCTTACTGTGCTGTTTCATATATATGGTGGTCTCA 2183
Qy 1021 ----- 1020
Db 2184 GTGTCAGTTGGTGAATAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCCCTAGC 2243
Qy 1021 ----- 1020
Db 2244 CTCTCTAGGTTATGTGGTTGTAGTAGACACAGGGGATCTCTGTCAACCGGGCTTAA 2303
Qy 1021 ----- 1020
```


Db 2304 ATTGGAAGGCGCTTTAAATATAAAATGGGTCAAAATAGAAATTCAGCATCAGTGAAGG 2363
Qy 1021 ----- 1020
Db 2364 AC7CCAATATCTAGCTTCTCGATATGATTTTCATTGACTTAGATCGTGTGGGCATCCACGG 2423
Qy 1021 ----- 1020
Db 2424 CTGGTCTATGGAGGATACCTCTCCCTGATGGCATTAAATGCAGAGGTTCAGATATCTTCAG 2483
Qy 1021 GGTTCGTATTTGCTGGGGCCCCAGTCAC7CTGTGGATCTTCTATGATACAGGATACACGGA 1080
Db 2484 GGTTCGTATTTGCTGGGGCCCCAGTCAC7CTGTGGATCTTCTATGATACAGGATACACGGA 2543
Qy 1081 ACCTTATATGGGTCACTCTGACCAAGAGTGAACAGGGCTATTACTTAGGATCTGTGGCCAT 1140
Db 2544 ACCTTATATGGGTCACTCTGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCCAT 2603
Qy 1141 GCAAGCAGAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGCTTCTCTGGA 1200
Db 2604 GCAAGCAGAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGCTTCTCTGGA 2663
Qy 1201 TGGAGATGTCATTTTGGACATACAGTATATTACTGAGTCTTTTACTGAGGGCTGAAA 1260
Db 2664 TGGAGATGTCATTTTGGACATACAGTATATTACTGAGTCTTTTACTGAGGGCTGAAA 2723
Qy 1261 GCCATATGATTTACAGATCTATCTCTCAGGAGAGACACAGATAAGAGTCTCTGAATCGGG 1320
Db 2724 GCCATATGATTTACAGATCTATCTCTCAGGAGAGACACAGATAAGAGTCTCTGAATCGGG 2783
Qy 1321 AGAACATTAATGAACTGATCTTTTGGACATACAGTATATTACTGAGTCTTTTACTGAGGGCTGAAA 1380
Db 2784 AGAACATTAATGAACTGATCTTTTGGACATACAGTATATTACTGAGTCTTTTACTGAGGGCTGAAA 2843
Qy 1381 TGCTCTAAAGTGCATATATTTTGGACCTGTGTGAGAACTCTCTGCTATACATGCTGCTATTT 1440
Db 2844 TGCTCTAAAGTGCATATATTTTGGACCTGTGTGAGAACTCTCTGCTATACATGCTGCTATTT 2903
Qy 1441 AACAATGAGGAGGTTTAAATCAACAGAGAAACACAGAAATGATCATCATCTTTTGATACC 1500
Db 2904 AACAATGAGGAGGTTTAAATCAACAGAGAAACACAGAAATGATCATCATCTTTTGATACC 2963
Qy 1501 TGCCATGTAACATCTACTCTCTGAAATAAATGTTGTCATGAGGGGCTACGGTTTGT 1560
Db 2964 TGCCATGTAACATCTACTCTCTGAAATAAATGTTGTCATGAGGGGCTACGGTTTGT 3023
Qy 1561 GGTAGTAATCTAATACCTTAAACCCACATGCTCAAAATCAAAATGATCATCATCTTTTGATACC 1620
Db 3024 GGTAGTAATCTAATACCTTAAACCCACATGCTCAAAATCAAAATGATCATCATCTTTTGATACC 3083
Qy 1621 GACCCAGCAATACCAAGAAATTAATAAAAAAAAAAAAA 1657
Db 3084 GACCCAGCAATACCAAGAAATTAATAAAAAAAAAAAAA 3120

RESULT 15
ABK12892
ID ABK12892 standard; cDNA; 3106 BP.
XX
AC ABK12892;
XX
DT 09-APR-2002 (first entry)
XX
DE Human protease PRTS-9 cDNA sequence.
KW Human; protease; PRTS; gastrointestinal; Crohn's disease; cancer;
KW cardiovascular; atherosclerosis; autoimmune disorder; dermatitis;
KW inflammatory disorder; acquired immunodeficiency syndrome; AIDS;
KW cell proliferative disorder; developmental disorder; epilepsy;
KW Duchenne muscular dystrophy; epithelial disorder; neurological disorder;
KW reproductive disorder; endometriosis; ss.
XX
OS Homo sapiens.

XX Key Location/Qualifiers
FH CDS 203..2851 ;
FT /*tag= a
FT /product= "Human protease PRTS-9"
XX
FN WO200198468-A2.
XX
PD 27-DEC-2001.
XX
PF 13-JUN-2001; 2001WO-US019178.
XX
PR 16-JUN-2000; 2000US-0212336P.
PR 22-JUN-2000; 2000US-0213955P.
PR 29-JUN-2000; 2000US-0215396P.
PR 07-JUL-2000; 2000US-0216821P.
PR 14-JUL-2000; 2000US-0218946P.
XX (INCY-) INCYTE GENOMICS INC.
PA Yue H, Elliott VS, Gandhi AR, Lal P, Au-Young J, Tribouley CM;
XX Deleage AM, Baughn MR, Nguyen DB, Lee EA, Hafalia A, Khan FA;
PI Walia NK, Yao MG, Lu DAM, Patterson C, Tang YT, Walsh RT;
PI Azimzai Y, Lu Y, Rankumar J, Xu Y, Reddy R, Das D, Kearney L;
PI Kallick DA;
XX
DR WP1; 2002-090437/12.
DR P-PSDB; AAU74749.
XX
XX Twenty one human proteases (referred to as PRTS-1 to PRTS-21), useful in
PT the diagnosis, treatment and prevention of gastrointestinal (e.g.
PT gastritis), cardiovascular (e.g. atherosclerosis) and cell proliferative
PT (e.g. cancer) disorders.
XX
PS Claim 5; Page 166-167; 177pp; English.
XX
CC The present invention relates to twenty one new human proteases, referred
CC to as PRTS-1 to PRTS-21. The PRTS polynucleotides and polypeptides of the
CC invention are useful in the diagnosis, treatment and prevention of
CC gastrointestinal e.g. gastritis, esophageal carcinoma and Crohn's
CC disease, cardiovascular e.g. atherosclerosis, hypertension and myocardial
CC infarction, autoimmune/inflammatory e.g. acquired immunodeficiency
CC syndrome (AIDS), allergies and osteoarthritis, cell proliferative e.g.
CC cancer, developmental e.g. Duchenne and Becker muscular dystrophy,
CC epithelial e.g. dermatitis, neurological e.g. epilepsy and Alzheimer's
CC disease and reproductive e.g. infertility and endometriosis disorders.
CC Numerous other examples of each disorder are given in the specification.
CC The present nucleic acid sequence encodes the human protease PRTS-9
CC protein of the invention
XX
SQ Sequence 3106 BP; 928 A; 633 C; 704 G; 841 T; 0 U; 0 Other;
Query Match 80.5%; Score 1344; DB 6; Length 3106;
Best Local Similarity 84.6%; Pred. No. 0;
Matches 1654; Conservative 0; Mismatches 0; Indels 300; Gaps 1;
Qy 1 AACAGGTACAGCAAAATCTTAAAGTCACTTTTAAAGATGTGAGAAATTAATGATGCTGA 60
Db 1153 AACAGGTACAGCAAAATCTTAAAGTCACTTTTAAAGATGTGAGAAATTAATGATGCTGA 1212
Qy 61 AGGAAGATCATAGATGTCTATAGATAGGACATTAATCAACCTTTTGAGATCTTATTTGA 120
Db 1213 AGGAAGATCATAGATGTCTATAGATAGGACATTAATCAACCTTTTGAGATCTTATTTGA 1272
Qy 121 AGGAGTTGAATATATTGCCAGAGCTGATGAGTCTCTGAGGAAATATGCTTGTGTCAT 180
Db 1273 AGGAGTTGAATATATTGCCAGAGCTGATGAGTCTCTGAGGAAATATGCTTGTGTCAT 1332
Qy 181 CCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGATATTATTTAT 240
Db 1333 CCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGATATTATTTAT 1392
Qy 241 CCCAGTGAAGATGATGTTATGGAAGGCAGAGACTCATTGAGTCAGTCTGATCTCTGT 300

```
Db 1393 CCCAGTAGAAGATGATGTTATGGAAGGCGAGAGACTCACTGAGTCAGTGCGTATTCCTGT 1452
Qy 301 GAGCCCACTAAATATCTATGAGAAACACACAGACATCTGGATAAATATCCATGACATCTT 360
Db 1453 GAGCCCACTAAATATCTATGAGAAACACACAGACATCTGGATAAATATCCATGACATCTT 1512
Qy 361 TCATGTTTTCCCAAGTCACGAGAGGAAATGAGTTTATTTTTGGCTCTGAATGCAA 420
Db 1513 TCATGTTTTCCCAAGTCACGAGAGGAAATGAGTTTATTTTTGGCTCTGAATGCAA 1572
Qy 421 AACAGGTTCCGTCATTTATACAAATTAACATTAATTTTAAAGGAAACCAAAATATAACG 480
Db 1573 AACAGGTTCCGTCATTTATACAAATTAACATTAATTTTAAAGGAAACCAAAATATAACG 1632
Qy 481 ATCCAGTGGTGGCTGCCCTGCCCTCAAGTGATTTCAAGTGCTCTATCAAGAGGAGATAGC 540
Db 1633 ATCCAGTGGTGGCTGCCCTGCCCTCAAGTGATTTCAAGTGCTCTATCAAGAGGAGATAGC 1692
Qy 541 AATTACAGTGGTGAATGGAAAGTCTTTGGCCGGCATGGATCTAATATCCAAGTTGATGA 600
Db 1693 AATTACAGTGGTGAATGGAAAGTCTTTGGCCGGCATGGATCTAATATCCAAGTTGATGA 1752
Qy 601 AGTCAGAAGGCTGGTATATTTTGAAGGCCACCAAGACTCCCTTTAGAGCATCACCTGTA 660
Db 1753 AGTCAGAAGGCTGGTATATTTTGAAGGCCACCAAGACTCCCTTTAGAGCATCACCTGTA 1812
Qy 661 CGTAGTCAGTTACGTAAATCCCTGGAGAGGTGACAGGCTGACTGACCGTGCTACTACA 720
Db 1813 CGTAGTCAGTTACGTAAATCCCTGGAGAGGTGACAGGCTGACTGACCGTGCTACTACA 1872
Qy 721 TTCTTCTCCTCAGTCAGCAGCTGACTCTTTTAAAGTAAGTATAGTAAACAGAGAA 780
Db 1873 TTCTTCTCCTCAGTCAGCAGCTGACTCTTTTAAAGTAAGTATAGTAAACAGAGAA 1932
Qy 781 TCACACTGTGTGTCCTTTTCAAGCTATCAAGTCTCGAAGTGAACCAACTTGCAGAAAC 840
Db 1933 TCACACTGTGTGTCCTTTTCAAGCTATCAAGTCTCGAAGTGAACCAACTTGCAGAAAC 1992
Qy 841 AAAGGAATTTGGGCCACCATTTTGGATTTCAGAGGTCCTTCTCTGACTATATCTCTCC 900
Db 1993 AAAGGAATTTGGGCCACCATTTTGGATTTCAGAGGTCCTTCTCTGACTATATCTCTCC 2052
Qy 901 AGAAATTTTCTCTTTTGAAGTACTACTGATTTACATTTGTATGGGATGCTCTACAGGC 960
Db 2053 AGAAATTTTCTCTTTTGAAGTACTACTGATTTACATTTGTATGGGATGCTCTACAGGC 2112
Qy 961 TCATGATCTACAGCCTGGAAAGAAATATCCTACTGTGCTGTTTATATATGTTGGTCTCTCA 1020
Db 2113 TCATGATCTACAGCCTGGAAAGAAATATCCTACTGTGCTGTTTATATATGTTGGTCTCTCA 2172
Qy 1021 ----- 1020
Db 2173 GGTGCAGTTGGTGAATAANTCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCCTAGC 2232
Qy 1021 ----- 1020
Db 2233 CTCTCTAGGTTATGTTGTTAGTGATAGACACAGGGGATCCTGTGTCAGGGGGCTTAA 2292
Qy 1021 ----- 1020
Db 2293 ATTTGAGGGCCTTTAAATATAAATAGGGTCAAAATAGAAATTTGACGATCAGGTGAAGG 2352
Qy 1021 ----- 1020
Db 2353 ACTCCAAATATCTAGCTTCTCGATATGATTTCAITTAGACTTAGATCGTGTGGGCATCCAGG 2412
Qy 1021 ----- 1020
Db 2413 CTGGTCCCTATGGAGGATACCTCTCCCTGATGGCAATTAATGACAGGTCAGATATCTTCAG 2472
Qy 1021 GGTGTGCTATGTGGGGCCCGAGTCACCTCTGTGGATCTTCTATGATACAGGATACACGGA 1080
```

```
Db 2473 GGTTGCTATTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTTATGATACAGGATACACGGA 2532
Qy 1081 ACGTTATATGGGTCAACCCCTGACCCAGATGAAACAGGGCTATTTACTTTAGGATCTGTGGCCAT 1140
Db 2533 ACGTTATATGGGTCAACCCCTGACCCAGATGAAACAGGGCTATTTACTTTAGGATCTGTGGCCAT 2592
Qy 1141 GCAAGCAGAAAAAGTTCCCTCTGAAACCAAAATCGTTTACTTGCTCTTTACATGTTTCTCTGGA 1200
Db 2593 GCAAGCAGAAAAAGTTCCCTCTGAAACCAAAATCGTTTACTTGCTCTTTACATGTTTCTCTGGA 2652
Qy 1201 TGGAATGTCCATTTTGGCACAACACAGTATATTTACTTGAGTTTTTTTGTAGTGAGGGCTGGA 1260
Db 2653 TGGAATGTCCATTTTGGCACAACACAGTATATTTACTTGAGTTTTTTTGTAGTGAGGGCTGGA 2712
Qy 1261 GCCATATGATTTACAGATCTATCCTCAGGAGAGACACAGCATAGAGTTCCCTGAATCGGG 1320
Db 2713 GCCATATGATTTACAGATCTATCCTCAGGAGAGACACAGCATAGAGTTCCCTGAATCGGG 2772
Qy 1321 AGAACATTATGAACCTGCATCTTTTGGCAGTACCTTCAAGAAAAACCTTGGATCACGTTTGC 1380
Db 2773 AGAACATTATGAACCTGCATCTTTTGGCAGTACCTTCAAGAAAAACCTTGGATCACGTTTGC 2832
Qy 1381 TGCTCTAAAAGTGATATAATTTTGGACCTGTGTAGAACTCTCTGGTATACACTGGCTATTT 1440
Db 2833 TGCTCTAAAAGTGATATAATTTTGGACCTGTGTAGAACTCTCTGGTATACACTGGCTATTT 2892
Qy 1441 AACCAATGAGGAGGTTTAAATCAACAGAAAAACACAGAAATTTGATCATCATTTTGTATACC 1500
Db 2893 AACCAATGAGGAGGTTTAAATCAACAGAAAAACACAGAAATTTGATCATCATTTTGTATACC 2952
Qy 1501 TGCCATGTAAACATCTACTCTCTGAAAAATAAATGTGTGTGCCATGCGAGGGTCTACCGTTTGT 1560
Db 2953 TGCCATGTAAACATCTACTCTCTGAAAAATAAATGTGTGTGCCATGCGAGGGGTCTACCGTTTGT 3012
Qy 1561 GGTAGTAATCTAATACCTTAAACCCACATGCTCAAAATCAAATGATACATATTCCTGAGA 1620
Db 3013 GGTAGTAATCTAATACCTTAAACCCACATGCTCAAAATCAAATGATACATATTCCTGAGA 3072
Qy 1621 GACCCAGCAATACCATAAGATTTACTAAAAAAA 1654
Db 3073 GACCCAGCAATACCATAAGATTTACTAAAAAAA 3106
```

Search completed: May 4, 2006, 09:09:25
Job time : 957.948 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 4, 2006, 02:59:51 ; Search time 6252.09 Seconds
(without alignments)
12489.839 Million cell updates/sec

Title: US-10-825-632-6
Perfect score: 1669
Sequence: 1 aacaggtagcaaatccta.....aaaaaaaaaaaaaaaaaaaaa 1669

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST: *
1: gb_est1: *
2: gb_est2: *
3: gb_est3: *
4: gb_hc: *
5: gb_est4: *
6: gb_est5: *
7: gb_est6: *
8: gb_est7: *
9: gb_gss1: *
10: gb_gss2: *
11: gb_gss3: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1568	93.9	2292	4 CR609512	full-length
2	904.6	54.2	3143	4 AK016546	Mus muscu
3	888.4	53.2	2649	10 AY411615	AY411615 Homo sapi
4	879.4	52.7	5517	4 AK029788	AK029788 Mus muscu
5	842.8	50.5	2649	10 AY411616	AY411616 Pan trogl
6	820.6	49.2	910	5 BX372276	BX372276 BX372276
7	804.4	48.2	902	8 DN517062	DN517062 1257572 M
8	730.6	43.8	857	8 DN519771	DN519771 1262448 M
9	688.4	41.2	853	2 BI223892	BI223892 602941035
10	675.6	40.5	2634	10 AY411617	AY411617 Mus muscu
11	672.6	40.3	1013	7 CN647232	CN647232 ILLUMIGEN
12	672.6	40.3	1027	7 CN646612	CN646612 ILLUMIGEN
13	660	39.5	993	2 BG259714	BG259714 602380072
14	646.4	38.7	1292	4 AF175225	AF175225 Homo sapi
15	644	38.6	722	5 BU631054	BU631054 UI-H-PEI-
16	640.8	38.4	804	5 BU687687	BU687687 UI-CF-ECL
17	637	38.2	1038	1 AL582206	AL582206 AL582206
18	628	37.6	634	1 AI819365	AI819365 wg61c02.x
19	623.6	37.4	760	1 AI917735	AI917735 tt1ld07.x
20	622.8	37.3	673	1 AW303607	AW303607 xv21b12.x
21	620	37.1	864	1 AL542617	AL542617 AL542617
22	607.6	36.4	1278	4 AF173382	AF173382 Homo sapi

23	601	36.0	1265	4 AF176779	AF176779 Homo sapi
24	599.6	35.9	727	6 CF537710	CF537710 UI-M-GIO-
25	596	35.7	596	6 CB153590	CB153590 K-EST0211
26	591.4	35.4	597	5 BU784521	BU784521 in16d12.x
27	588.4	35.3	692	8 DN514951	DN514951 1252126 M
28	578	34.6	753	1 AL040398	AL040398 DKFZP434A
29	574.4	34.4	577	1 AI243857	AI243857 qn89f09.x
30	556.2	33.3	1021	2 BG974587	BG974587 602844820
31	556	33.3	1041	2 BI084090	BI084090 602869453
32	555.4	33.3	957	5 BQ675006	BQ675006 AGENCOURT
33	553.4	33.2	627	2 BG752328	BG752328 602730802
34	553	33.1	968	5 BQ671635	BQ671635 AGENCOURT
35	552	33.1	910	5 BQ675260	BQ675260 AGENCOURT
36	551.4	33.0	590	3 BI711119	BI711119 id90908.x
37	549	32.9	639	1 AW958665	AW958665 EST370735
38	547	32.8	897	5 BU191638	BU191638 AGENCOURT
39	537	32.2	671	8 CX787143	CX787143 HESC3 81
40	535	32.1	762	6 CD352521	CD352521 UI-M-GIO-
41	532.6	31.9	732	6 CF744232	CF744232 UI-M-GVO-
42	529.4	31.7	536	3 BI710784	BI710784 id90908.y
43	528.6	31.7	774	6 CB233750	CB233750 AGENCOURT
44	524	31.4	735	1 AL043338	AL043338 DKFZP434O
45	521.2	31.2	719	8 DR768184	DR768184 ILLUMIGEN

ALIGNMENTS

RESULT 1
LOCUS CR609512 2292 bp mRNA linear HTC 21-JUL-2004
DEFINITION full-length cDNA clone CSODL005YD02 of B cells (Ramos cell line)
Cot 25-normalized of Homo sapiens (human).
ACCESSION CR609512
VERSION CR609512.1 GI:50490319
KEYWORDS HTC; CNSLT_cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 2292)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
REFERENCE 2 (bases 1 to 2292)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (20-JUL-2004) Genoscope - Centre National de Sequençage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr) :
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES Location/Qualifiers
source 1..2292
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clones="CSODL005YD02"
/tissue_type="B cells (Ramos cell line) Cot 25-normalized"
/plasmid="pCMVSPORT_6"

ORIGIN

Query Match 93.9%; Score 1568; DB 4; Length 2292;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1590; Conservative 0; Mismatches 0; Indels 12; Gaps 1;

QY 1 AACAGGTACAGCAAAATCCCTAAAGTCACTTTTAAAGTGTACAGAAATATATGATGATGCTGA 60
Db 703 AACAGGTACAGCAAAATCCCTAAAGTCACTTTTAAAGTGTACAGAAATATATGATGATGCTGA 762
QY 61 AGGAAGGATCATAGATGTCTATAGATAAGGAACTAAATCAACCTTTTGAGATTCCTATTGA 120
Db 763 AGGAAGGATCATAGATGTCTATAGATAAGGAACTAAATCAACCTTTTGAGATTCCTATTGA 822
QY 121 AGGAGTTGAATATATATGCGCAGAGCTGGATGGACTCCTGAGGGAATAATGCTGGTCCAT 180
Db 823 AGGAGTTGAATATATATGCGCAGAGCTGGATGGACTCCTGAGGGAATAATGCTGGTCCAT 882
QY 181 CCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGTCTCACCTGAATATATTTAT 240
Db 883 CCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGTCTCACCTGAATATATTTAT 942
QY 241 CCCAGTAGAAGATGATGTTATGGAAGGAGGAGAGACTCAATGAGTCAGTGCCTGATTCCTGT 300
Db 943 CCCAGTAGAAGATGATGTTATGGAAGGAGGAGAGACTCAATGAGTCAGTGCCTGATTCCTGT 1002
QY 301 GAGCCCACTAATATCTATGAAGAACACACACATCTGGATAATATCCATGACATCTT 360
Db 1003 GAGCCCACTAATATCTATGAAGAACACACACATCTGGATAATATCCATGACATCTT 1062
QY 361 TCATGTTTTTCCCAAGTCCAGAGGAGAAATTCAGTTTATTTTGGCTCTGGAATGCAA 420
Db 1063 TCATGTTTTTCCCAAGTCCAGAGGAGAAATTCAGTTTATTTTGGCTCTGGAATGCAA 1122
QY 421 AACAGGTTTCGCTCATTTATACAAATTAATCATCTATTTTAAAGGAAAGCAATATAACG 480
Db 1123 AACAGGTTTCGCTCATTTATACAAATTAATCATCTATTTTAAAGGAAAGCAATATAACG 1182
QY 481 ATCCAGTGGTGGGCTGCTGCTCCAAAGTATTCAGTGTCTTATCAAGTGCTCTATCAAGGAGGATAGC 540
Db 1183 ATCCAGTGGTGGGCTGCTGCTCCAAAGTATTCAGTGTCTTATCAAGTGCTCTATCAAGGAGGATAGC 1242
QY 541 AATTACCAAGTGGTGAATGGGAAGTCTTGGCCGGATGATCTAATATCCAAAGTTGATGA 600
Db 1243 AATTACCAAGTGGTGAATGGGAAGTCTTGGCCGGATGATCTAATATCCAAAGTTGATGA 1302
QY 601 AGTCAGAAAGGCTGGTATATTTTGAAGGCCACCAAGACTCCCTTTAGAGCATCACCTGTA 660
Db 1303 AGTCAGAAAGGCTGGTATATTTTGAAGGCCACCAAGACTCCCTTTAGAGCATCACCTGTA 1362
QY 661 CGTAGTCAGTTACGTAAATCCTGGAGAGGTGACAAAGGCTGACTGACCGTGCTACTCACA 720
Db 1363 CGTAGTCAGTTACGTAAATCCTGGAGAGGTGACAAAGGCTGACTGACCGTGCTACTCACA 1422
QY 721 TTCTTGCTGCATCAGTCAGCACTGCTACTTCTTTATAGTAAAGTATAGTACCAAGAGAA 780
Db 1423 TTCTTGCTGCATCAGTCAGCACTGCTACTTCTTTATAGTAAAGTATAGTACCAAGAGAA 1482
QY 781 TCCACACTGTGTCCCTTTACAAAGCTATCAAGTCTGGAAGTACCCAACTTCCAAAC 840
Db 1483 TCCACACTGTGTCCCTTTACAAAGCTATCAAGTCTGGAAGTACCCAACTTCCAAAC 1542
QY 841 AAAGGAATTTGGCCACCAATTTGGATTCAGCAGGCTCTCTTCTGACTATATCTCTCC 900
Db 1543 AAAGGAATTTGGCCACCAATTTGGATTCAGCAGGCTCTCTTCTGACTATATCTCTCC 1602
QY 901 AGAAATTTCTTTTGAAGTACTACTGATTTACATTTGATGGATGCTCTCAAGCC 960
Db 1603 AGAAATTTCTTTTGAAGTACTACTGATTTACATTTGATGGATGCTCTCAAGCC 1662
QY 961 TCATGATCTACAGCTGGAAAGAAATATCTACTGTGTGCTTTCATATATGTTGGTCTCA 1020
Db 1663 TCATGATCTACAGCTGGAAAGAAATATCTACTGTGTGCTTTCATATATGTTGGTCTCA 1722
QY 1021 GGTGTCTATTGCTGGGGCCCACTCACTCTGTGGATCTTCTATGATACAGGATACACGGA 1080
Db 1723 GGTGTCTATTGCTGGGGCCCACTCACTCTGTGGATCTTCTATGATACAGGATACACGGA 1782
QY 1081 ACGTTATATGGGTCACCCCTGACCAGAAATGAACAGGGCTATTACTTAGGATCTGTGGCCAT 1140

Db 1783 ACGTTATATGGGTCACCCCTGACCGAATGAACAGGGCTATTACTTAGGATCTGTGGCCAT 1842
QY 1141 GCAAGCAGAAAAGTTCCTCTCTGAAACCAATCGTTTACTGCTCTTACATGTTTCTCTGGA 1200
Db 1843 GCAAGCAGAAAAGTTCCTCTCTGAAACCAATCGTTTACTGCTCTTACATGTTTCTCTGGA 1902
QY 1201 TGAGAATGTCATTTTGGACATACCAAGTATATATCTAGTGTGAGGGCTGGAAA 1260
Db 1903 TGAGAATGTCATTTTGGACATACCAAGTATATATCTAGTGTGAGGGCTGGAAA 1962
QY 1261 GCCATATGATTTTACAGATCTATCCTCAGGAGAGACACAGATTAAGAGTTCCTGAATCGG 1320
Db 1963 GCCATATGATTTA-----CAGGAGAGACACAGATTAAGAGTTCCTGAATCGG 2010
QY 1321 AGAACATTTAGAACTGCTATTTTGGCACTACCTTCAAGAAAAACCTTGGATCACGTATTGC 1380
Db 2011 AGAACATTTAGAACTGCTATTTTGGCACTACCTTCAAGAAAAACCTTGGATCACGTATTGC 2070
QY 1381 TGCTCTAAAAGTGTATATAATTTTGGCTGTGTAGAACTCTCTGGTATACACTGGCTATTT 1440
Db 2071 TGCTCTAAAAGTGTATATAATTTTGGCTGTGTAGAACTCTCTGGTATACACTGGCTATTT 2130
QY 1441 AACCAATGAGGAGTTTAAATCAACAGAAAACACAGAAATGATCATCATTTTGAATACC 1500
Db 2131 AACCAATGAGGAGTTTAAATCAACAGAAAACACAGAAATGATCATCATTTTGAATACC 2190
QY 1501 TGCCATGTAACATCTACTCTGAAATATACTGTTGGTCCATGCAGGGCTTACGGTTGT 1560
Db 2191 TGCCATGTAACATCTACTCTGAAATATACTGTTGGTCCATGCAGGGCTTACGGTTGT 2250
QY 1561 GGTAGTATCTATATACCTTAAACCCACATGCTCAAAATCAA 1602
Db 2251 GGTAGTATCTATATACCTTAAACCCACATGCTCAAAATCAA 2292

RESULT 2
AKO16546
LOCUS
DEFINITION
3143 bp mRNA linear HTC 03-APR-2004
Mus musculus adult male testis cDNA, RIKEN full-length enriched library, clone:4932434F09 product:D1PEPTIDYL PEPTIDASE 8 homolog [Homo sapiens], full insert sequence.
AKO16546
VERSION
HTC; CAP trapper.
KEYWORDS
Mus musculus (house mouse)
SOURCE
Mus musculus
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
10349636
AUTHORS
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
JOURNAL
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
11042159
PUBMED
2
AUTHORS
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishimura, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, I., Izawa, M., Ohara, E., Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, I., Matsura, S., Kawa, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
Riken integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
3
AUTHORS
TITLE
JOURNAL
PUBMED
REFERENCE

Qy	841	AAAGGAATTTTGGGCGACCAATTTTGGATTACAGAGTCTCTTCTCTGACTATATCTCTCC	900
Db	2147	AAAGGAATTTTGGGCGACCAATTTTGGATTACAGAGTCTCTTCTCTGACTATATCTCTCC	2206
Qy	901	AGAAATTTTCTCTTTTGAAGTACTACTGATTTACATGTATGGATGCTCTACAAAGCC	960
Db	2207	AGAAATTTTCTCTTTTGAAGTACTACTGATTTACATGTATGGATGCTCTACAAAGCC	2266
Qy	961	TCATGATCTACAGCCCTGGAAAGAAATATCTCTACTGTCTGTTCATATATATGGTCTCTCA	1020
Db	2267	TCATGACCTACAAACCTGGAAAGAAATATCTCTACTGTCTGTTCATATATATGGTCTCTCA	2326
Qy	1021	-----	1020
Db	2327	GGTGACGTGGTGAAACAATCGGTTTAAAGGAGTCAAGTATTTCCGCGTGAACACCCCTGGC	2386
Qy	1021	-----	1020
Db	2387	CTCCCTGGGTTATGTGGTTGTGGTGATAGACAACAGGGGATCCTGTCCCGAGGACTTAA	2446
Qy	1021	-----	1020
Db	2447	ATTGAAGGCGCCTTTAAATATAAAATGGGTCAATAGAAATCGATCAAGTGAAGG	2506
Qy	1021	-----	1020
Db	2507	ACTCCAGTACTAGCATCTCAGTATGACTTCAATTGGATCGAGTGGGCATCCACGG	2566
Qy	1021	-----	1020
Db	2567	CTGGTCTATGGTGGTACCTCTCCCTGATGGCATTAATGCAGAGGTCCGATCTTCGG	2626
Qy	1021	GGTGTCTATTTCTGGGCGCCAGTCACTCTGTGATCTTCTATGATACAGGATACAGGA	1080
Db	2627	GGTGTCTATTTCTGGGCGCCAGTCACTCTGTGATCTTCTATGATACAGGATACAGGA	2686
Qy	1081	ACGTTATATGGTCAACCTGACACAGATGAACAGGGCTATTAATTAGGATCTGTGGCCAT	1140
Db	2687	GCCTATATGGGTCAACCTGACACAGATGAACAGGGCTATTAATTAGGATCTGTGGCCAT	2746
Qy	1141	GCAGCAGAAAAGTTCCCTCTGTAACCAAAATCGTTTACTCTCTTACATGGTTCTCTGA	1200
Db	2747	GCAGCGGAGAAAGTTCCCTCTGTAACCAAAATCGTTTACTCTCTTACATGGTTCTCTGA	2806
Qy	1201	TGAGATGTCCATTTTGACATACAGTATATTAAGTATTTTCTAGTGGGCTGGAAA	1260
Db	2807	TGAGATGTTCATTTTGACATACAGTATATTAAGTATTTTCTAGTGGGCTGGAAA	2866
Qy	1261	GCATATGATTTTACAGATCTATCTCTCAGGAGAGACAGCATAAAGATTCCTGAAATCGGG	1320
Db	2867	GCATATGATTTTACAGATCTATCTCTCAGGAGAGACAGCATAAAGATTCCTGAAATCGGG	2926
Qy	1321	AGAACATTTATGAATGCACTCTTTTGCACTACCTTTCAAGAAACCTTTGGATCAGTATTCG	1380
Db	2927	AGAACATTTATGAATGCACTCTTTTGCACTACCTTTCAAGAAACCTTTGGATCAGTATTCG	2986
Qy	1381	TGCTCTAAAAGTGATATAATTTTACCTGTGTAGACTCTCTGTGATACACATGGCTATT	1440
Db	2987	TGCTCTAAAAGTGATATAATTTTACCTGTGTAGACTCTCTGTGATACACATGGCTATT	3039
Qy	1441	BAACAAATGAGGAGTTTAACTACAGAAACACAGATTTGATCATCAATTTGATACC	1500
Db	3040	BAACAAACAAAGAGGGTTAGCCATGGAGAGCCAGAGTTGATGAACCTTCTTTGTACC	3099
Qy	1501	TGCCATGTAACTCTCTCTGAAATAAATGTGGTGCCATGC	1543
Db	3100	TGCCCATNAGATCTATTCTTGAAGATTAATTTGGTGCCATGC	3142

RESULT 3
AY411615
LOCUS

AY411615

2649 bp

DNA

linear

GSS 16-DEC-2003

DEFINITION	Homo sapiens DPP8 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.
ACCESSION	AY411615
VERSION	AY411615.1 GI:39767583
KEYWORDS	GSS.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	1 (bases 1 to 2649) Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarimal, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Shinsky, J.J., Adams, M.D. and Cargill, M. Infering nonneutral evolution from human-chimp-mouse orthologous gene trios
TITLE	Science 302 (5652), 1960-1963 (2003)
JOURNAL	14671302
PUBMED	2 (bases 1 to 2649)
REFERENCE	Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarimal, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Shinsky, J.J., Adams, M.D. and Cargill, M. Direct Submission Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA This sequence was made by sequencing genomic exons and ordering them based on alignment.
FEATURES	Location/Qualifiers 1..2649 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" <1..>2649 /gene="DPP8" /locus_tag="HCM4283"
ORIGIN	Query Match 53.2%; Score 888.4; DB 10; Length 2649; Best Local Similarity 70.6%; Pred. No. 1.5e-218; Matches 1199; Conservative 0; Mismatches 200; Indels 300; Gaps 1;
Qy	1 AACAGGTACAGCAATCTTAAAGTCACCTTTTAAAGTCTCAGAAATATGATGATCTGA 60
Db	951 AACAGGTACAGCAATCTTAAAGTCACCTTTTAAAGTCTCAGAAATATGATGATGCTGA 1010
Qy	61 AGGAAGGATCATAGATGTCATAGATAAGGAACTTAATCAACCTTTTGAGATTCCTATTGA 120
Db	1011 AGGAAGGATTCATATATGTCAGAGCTGGATGCTCTGAGGAAATATGCTTTGGTCCAT 1070
Qy	121 AGGATTTGATATATGTCAGAGCTGGATGCTCTGAGGAAATATGCTTTGGTCCAT 180
Db	1071 NNN 1130
Qy	181 CCTACTAGATCGCTCCAGACTCCGCTACAGATAGTGTTCATCTCACCTGAATTTATTTAT 240
Db	1131 CCTACTAGATCGCTCCAGACTCCGCTACAGATAGTGTTCATCTCACCTGAATTTATTTAT 1190
Qy	241 CCCAGTAGAAGATGATGTTATGGAAGGAGAGAGCTCATTGAGTCTGAGTCTGATCTCTGT 300
Db	1191 CCCAGTAGAAGATGATGTTATGGAAGGAGAGAGCTCATTGAGTCTGAGTCTGATCTCTGT 1250
Qy	301 GAGCCACTAATTTATCTATGAAGAAACCAACAGACATCTGGATAAATATCCATGACATCTT 360
Db	1251 GAGCCACTAATTTATCTATGAAGAAACCAACAGACATCTGGATAAATATCCATGACATCTT 1310
Qy	361 TCATGTTTTCCTCCAAAGTCACGAAGAGGAAATTTAGTTTATTTTTCCTCTGAAATGCAA 420
Db	1311 TCATGTTTTCCTCCAAAGTCACGAAGAGGAAATTTAGTTTATTTTTCCTCTGAAATGCAA 1370
Qy	421 AACAGGTTTCCTCATTTATATAAAATTTACATCTATTTTAAAGGAAGCAAAATAAAGC 480

Db 1371 AACAGGTTCCGTCATTTATACAAATTAATCATCTATTTTAAAGGAAGCAATATAAAG 1430
Qy 481 ATCCAGTGGTGGCTGCTGCTCCCAAGTGAITTCGAAGTGCTCTATCAAGAGGAGATAGC 540
Db 1431 ATCCAGTGGTGGCTGCTGCTCCAAANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1490
Qy 541 AATTAACAAGTGGTGAATGGAGAGTTCCTGGCCGGCATGGATCTAATATCCAAAGTTGATGA 600
Db 1491 NN 1550
Qy 601 AGTCAGAAGCTGGTATATTTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTA 660
Db 1551 AGTCAGAAGCTGGTATATTTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTA 1610
Qy 661 CGTAGTCAGTTACGTTAAATCCTGGAGAGGTGAACAAGGCTGACTGACCGTGCTACTCACA 720
Db 1611 CGTAGTCAGTTACGTTAAATCCTGGAGAGGTGAACAAGGCTGACTGACCGTGCTACTCACA 1670
Qy 721 TTTCTGTGTCAGTCAGTCAGGACTGTGACTTCTTTTAAAGTAAAGTATAGTAAACGAGAA 780
Db 1671 TTTCTGTGTCAGTCAGTCAGGACTGTGACTTCTTTTAAAGTAAAGTATAGTAAACGAGAA 1730
Qy 781 TCCACACTGTGTGCTCCTTTTACAAGCTATCAAGTCTCTGAAGTACCCCAACTTCGAAAC 840
Db 1731 TCCACACTGTGTGCTCCTTTTACAAGCTATCAAGTCTCTGAAGTACCCCAACTTCGAAAC 1790
Qy 841 AAAGGAATTTTGGGCGCACCATTTTGGATTGACAGGCTCTCTCTGACTATATCTCTCC 900
Db 1791 AAAGGAATTTTGGGCGCACCATTTTGGATTGACAGGCTCTCTCTGACTATATCTCTCC 1850
Qy 901 AGAAATTTCTTTTGAAGTACTACTGGAATTAATGATGATGGATGCTCTCAAGCC 960
Db 1851 AGAAATTTCTTTTGAAGTACTACTGGAATTAATGATGATGGATGCTCTCAAGCC 1910
Qy 961 TCATGATCTACAGCTGGAAGAAATATCTACTGCTGTGTTTCATATATGTTGTCCTCA 1020
Db 1911 TCATGATCTACAGCTGGAAGAAATATCTACTGCTGTGTTTCATATATGTTGTCCTCA 1970
Qy 1021 ----- 1020
Db 1971 GGTGACGTTGGTGAATATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGATACCTTAGC 2030
Qy 1021 ----- 1020
Db 2031 CTCTCTAGGTTATGTTGGTTAGTATAGACAAACAGGGGATCTGTCAACCGAGGCTTAA 2090
Qy 1021 ----- 1020
Db 2091 ATTGAAGCGCCTTTAAATATATAAATGGTCAAAATAGAAATTCAGCATCAGGTGGAGG 2150
Qy 1021 ----- 1020
Db 2151 ACTCAATATCTAGCTTCTCGATATGATTTTCAATGACTTAGATCGTGGGCATCCAGG 2210
Qy 1021 ----- 1020
Db 2211 CTGTCTCTATGGAGGATACCTCTCCCTGATGGCATTAAATGCAGAGGTGAGATATCTTCAG 2270
Qy 1021 GGTGCTATTTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATCAGATACAGGA 1080
Db 2271 GGTGCTATTTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATCAGATACAGGA 2330
Qy 1081 ACGTTATATGGTCAACCTGACAGCAATCAACAGGCTATTACTTAGATCTGTGGCCAT 1140
Db 2331 ACGTTATATGGTCAACCTGACAGCAATCAACAGGCTATTACTTAGATCTGTGGCCAT 2390
Qy 1141 GCAAGCAGAAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGTTTCTCGGA 1200
Db 2391 GCAAGCAGAAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGTTTCTCGGA 2450
Qy 1201 TGAGATGTCATTTTGGACATACAGATATATCTAGATGTTTTTGTAGGGCTGAAA 1260
Db 2451 TGAGATGTCATTTTGGACATACAGATATATCTAGATGTTTTTGTAGGGCTGAAA 2510

Qy 1261 GCCATATGATTTACAGATCTATCTCTCAGGAGAGACACAGCATAAAGTTCCTGAATCGG 1320
Db 2511 GCCATATGATTTACAGATCTATCTCTCAGGAGAGACACAGCATAAAGTTCCTGAATCTGG 2570
Qy 1321 AGACATATTAACATGATCTTTTTCGACTACCTTCAAGAAAAACCTTGGATCAGCTATTGC 1380
Db 2571 AGACATATTAACATGATCTTTTTCGACTACCTTCAAGAAAAACCTTGGATCAGCTATTGC 2630
Qy 1381 TGCTCTAAAAGTGATATAA 1399
Db 2631 TGCTCTAAAAGTGATATAA 2649
RESULT 4
AK029788
LOCUS
DEFINITION
Mus musculus adult male testis cDNA, linear HTC 03-APR-2004
library, clone:4930560C15 product:D1PEPTIDYL PEPTIDASE 8 homolog
[Homo sapiens], full insert sequence.
AK029788
AK029788.1 GI:26081520
HTC; CAP trapper.
MUS musculus (house mouse)
SOURCE
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
10349636
2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
11042159
3
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Kono, H., Akiyama, J., Nishi, K., Kusunagi, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yanamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
11076861
4
The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 5517)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Kono, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, K., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,

Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, P., Takaku-Akahira, S.,
 Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
 Muramatsu, M. and Hayashizaki, Y.
 Direct Submission
 Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
 Physical and Chemical Research (RIKEN) Laboratory for Genome
 Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
 RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
 Kanagawa 230-0045, Japan. (E-mail: genome-res@gs.c.riken.jp,
 URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,
 Fax: 81-45-503-9216)
 COMMENT
 Library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.
 Division of Experimental Animal Research in Riken contributed to
 prepare mouse tissues.
 Please visit our web site for further details.
 URL: http://genome.gsc.riken.jp/
 URL: http://fantom.gsc.riken.jp/
 Location/Qualifiers
 1. 5517
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="FANTOM DB:4930560C15"
 /db_xref="taxon:10090"
 /clone="4930560C15"
 /sex="male"
 /tissue_type="testis"
 /clone_lib="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="adult"
 2221..4837
 misc_feature
 (note="DIPEPTIDYL PEPTIDASE 8 homolog (Homo sapiens)
 (SPTR109HMS, evidence: PASTY, 95.7%ID, 100%length,
 match=2649)
 putative")
 ORIGIN
 Query Match 52.7%; Score 879.4; DB 4; Length 5517;
 Best Local Similarity 72.6%; Pred. No. 3.8e-216;
 Matches 1362; Conservative 0; Mismatches 201; Indels 312; Gaps 5;
 QY 67 GATCATAGATGTCATAGATAAGAACTAATTCACCTTTTGAGATTCATTGAAGAGT 126
 DB 3205 GATTATAGATGTCATAGATAAGAACTGTTTCAACCTTTTCGAGATTCCTGTTGAGGAGT 3264
 QY 127 TGAATATATTGCCAGAGCTGGATGGACTCTCGAGGGAATAATATGCTTGGTCCATCTTACT 186
 DB 3265 TGAATATATTGCCAGAGCTGGATGGACTCTCGAGGGAATAATATGCTTGGTCCATCTTACT 3324
 QY 187 AGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCAGCTGAAATTAATTTATCCAGT 246
 DB 3325 AGACCGTTCCAGACTCACCTGCAGATAGTTCGATCTCCCTCAGTTATTCATCCAGT 3384
 QY 247 AGAAGATGATGTTATGGAAGGCAGAGACTCAATGAGTCAGTGCCTGATCTGTGAGCC 306
 DB 3385 AGAAGATGATGCCATGGACAGACAGACTTATAGAGTCAGTTCCTGACTCTGTGACCC 3444
 QY 307 ACTAATATCTATGAGAACACAGACATCTGGTAATATTCATCAGATCTTTCATGT 366
 DB 3445 ACTGATCATCTATGAGAACACAGACATCTGGTAATATTCATCAGATCTTTCATGT 3504
 QY 367 TTTTCCCAAGTCCAGAGAGAAATGAGTTATTTTTCCTCTGAGTGCAGAAACAGG 426
 DB 3505 TTTTCTCAACCTCATGAAAGATGAATTTGAGTTATTTTGGCTCTGAATGCAAAACAGG 3564
 QY 427 TTTCCGTATTTATACAAATATACATCTATTTTAAAGGAAGCAATATATAAGATCCAG 486
 DB 3565 TTTTCTGATCTGTATTAATACATCTATTTTAAAGGAGAGCAATATATAAGCGTCCAG 3624
 QY 487 TGTGGGCTCGCTGCTCCAGTGATTTCAAGTGTCTTCAAGAGGAGATGACAAATTAC 546
 DB 3625 TGTGGACTACCTGCCCCAAGTGATTTCAAGTGTCTTCAAGAGGAGCAATATAAGATTTAC 3684

QY 547 CAGTGTGAATGGGAAGTCTCTGGCCGCAATGGATCTATATATCCAAAGTTGATGAGTCAAG 606
 DB 3685 CAGTGTGAATGGGAAGTCTCTGGCCGCAATGGATCTATATATCTGGTTTGAAGCCAG 3744
 QY 607 AAGCTGGTATATTTGAAGGACCAAGACTCCCTTTAGAGCATCACCTGTAGTGTAGT 666
 DB 3745 AAGCTGGTGTACTTTGAAGGACCAAGACTCTCTTTGGAGCATCACCTGTAGTGTAGT 3804
 QY 667 CAGTTACGTAAATCTCTGGAGAGGTGACAAGGCTGACTGACCGTGGTCTACTCATATCTTTG 726
 DB 3805 CAGTTATGCAAAACCTCTGGAGAGTGTGAGGCTGACTGACCGTGGTCTACTCATACCTCTG 3864
 QY 727 CTGCATCAGTCAGACACTGTGACTCTTTTATAAGTAAAGTATAGTAAACCAAGAAATCCACA 786
 DB 3865 CTGCCTCAGCCGCGCATTTGTGACTTCTTCAATAGTAAAGTACAGCAACCAAGAAATCCACA 3924
 QY 787 CTGTGTCTCCCTTTCAAAAGCTATCAAGTCTCTCAAGATGACCCCAACTTGCACAAACAAAGGA 846
 DB 3925 CTGTGTCTCCCTCTCAAAACTCTCAAGTCTCTGAGGATGACCCAGTTCATAAAACAAAGGA 3984
 QY 847 ATTTTGGGCCACCAATTTTGGATTACAGAGTCTCTCTCTGCTGACTATATCTCTCTCCAGAAAT 906
 DB 3985 ATTTTGGGCCACCAATTTTGGATTACAGAGTCTCTCTCTGCTGACTACACCCCTCCAGAAAT 4044
 QY 907 TTTCTCTTTTGAAGTACTACTGGATTTTACATTTGTATGGATGCTCTACAAGCCCTCATGA 966
 DB 4045 TTTTCTTTTGAAGTACTACTGGATTTTACATCTGTATGGAATGTTGTATAGCCCTCATGA 4104
 QY 967 TCTACAGCTCTGGAAGAAATATCTTACTGCTGTGCTGCTATATATATGCTGCTCAGGTT-- 1024
 DB 4105 CTTACAACTCTGGAAGAAATATACCCCACTGTGTTATTCATATATGTTGTTGCCAGGTGCA 4164
 QY 1025 ----- 1024
 DB 4165 GCTGGTGAACAATCGGGTTTAAAGAGGTCAAGTATTTTCCGCTGGAACACCCCTGGGCTCCCT 4224
 QY 1025 ----- 1024
 DB 4225 GGGTTATGTGGTTGTGGTGTATAGACAACAGGGGATCTCTGTCCAGGAGACTTAAATTTGA 4284
 QY 1025 ----- 1024
 DB 4285 AGCGCCCTTTAAATATATAAAATGGGTCAATAGAAAATCGATGATCAAGTGGAGGACTCCA 4344
 QY 1025 ----- 1024
 DB 4345 GTACCTAGCATCTCAGTATGACTTCAATGACTTGGATCGAGTGGGCATCCACGGCTGGTC 4404
 QY 1025 -----GC 1026
 DB 4405 CTATGGTGGCTACCTCTCCCTGATGGCAATTAATGCAAGGTGCGGATATCTTCCGGGTGGC 4464
 QY 1027 TATTGCTGGGGCCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGACCGTTA 1086
 DB 4465 TATTGCTGGGGCCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAGCGCTA 4524
 QY 1087 TATGGTTCACCTGACAGCAAGTGAACAGGCTATTACTTAGGATCTGTGGCATGCAAGC 1146
 DB 4525 TATGGTTCACCTGACAGCAAGTGAACAGGCTATTACTTAGGATCTGTGGCATGCAAGC 4584
 QY 1147 AGAAAAGTTCCCTCTGAAACCAAAATGCTTTACTGCTCTTACATGTTTCTTGGATGAGAA 1206
 DB 4585 GGAAGATTTCCCTCTGAAACCAAAACCGGTTACTCTCTTGTGATGGTTCTTGGATGAGAA 4644
 QY 1207 TGTCCATTTTGCACATACACAGTATATTAATCTGAGTTTTTTAGTGAGGCTGGAAGGCCATA 1266
 DB 4645 TGTTCATTTTGCACACACAGTATATTTGCTGAGTTTTTTAGTGAGGCTGGAAGGCCATA 4704
 QY 1267 TGATTTTACAGATCTATCTCTCAGGAGAGACACAGCATTAAGAGTTCTGTAATCGGAGAACCA 1326
 DB 4705 TGACTTACAGATCTATCTCTCAGGAGAGGACACAGCATGAGTTCTGAGTCTGAGGAACA 4764


```
QY 1327 TTATGAACGATCTTTTGGACCTTACCTTCAAGAAACCTTGGATCACGTATTGCTGCTCT 1386
Db |||||||
Db 4765 CTATGAACCTACCTGCTCCACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTT 4824
QY 1387 AAAAGTGATATAATTTTGGACCTGTGAGAACTCTCTGGGTATACACTGGCTATTAAACAA 1446
Db |||||||
Db 4825 GAAAGTGATATACTCTGACGATGTAGAACT---TGGGTACACTG---TCTAACCA 4877
QY 1447 ATGAGGAGGTTTAAATCAACAGAAACACAGAAATGATCATCATCTTTTGTATACCTGCCAT 1506
Db |||||||
Db 4878 ACAAGAGGCTTAGCCAATAGAGAGCCAGAGTGTGATGAACCTACTTTTGTACCTGCCAC 4937
QY 1507 GTAACATCTACTCTCTGAAATAAATGTGTGTCATGCGAGGGTCTACGGTTTGG-TAG 1565
Db |||||||
Db 4938 ATAAGATCTATTCTTGAAGTAAATTTGGTGCATGCGAGAAATCTGCAAGTTTCATGSGATAG 4997
QY 1566 TAATCTTAATACCTTAACCCACATGTCACAAATCAAAATGATACATATTCTCTGAGAGACCC 1625
Db |||||||
Db 4998 TACTCTTAATACCTTAC----CATGCACATAAACTGATGACATGTTCCCAAGGATTG 5053
QY 1626 AGCAATACCATAGA 1640
Db |||||||
Db 5054 AACAGTACCTGAGGA 5068

RESULT 5
AY411616 2649 bp DNA linear GSS 16-DEC-2003
LOCUS Pan troglodytes Dpp8 gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION genomic survey sequence.
ACCESSION AY411616
VERSION AY411616.1 GI:39767584
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Pan.
REFERENCE 1 (bases 1 to 2649)
AUTHORS Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejarimal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Inferred nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 2649)
AUTHORS Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejarimal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source 1..2649
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
<1..>2649
/genes="Dpp8"
/locus_tag="HCM4283"
ORIGIN
Query Match 50.5%; Score 842.8; DB 10; Length 2649;
Best Local Similarity 68.1%; Pred. No. 9.6e-207;
Matches 1157; Conservative 0; Mismatches 242; Indels 300; Gaps 1;

QY 1 AACAGGTACAGCAAACTCTTAAGTCACTTTTAAGATGTCAGAAATAATGATGATGCTGA 60
|||||
```

```
Db 951 AACAGGTACAGCAAACTCTTAAGTCACTTTTAAGATGTCAGAAATAATGATGATGCTGA 1010
QY 61 AGGAAGAGTACATAGATGTCATAGATAAGGAACATAATCAACCTTTTGAGATTTCTATTGA 120
|||||
Db 1011 AGGAAGGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1070
QY 121 AGGAGTTGAATATATTGCCAGAGCTGGATGGACTCTGAGGGAATAATATGCTTGGTCCAT 180
|||||
Db 1071 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1130
QY 181 CCTACTAGATCGCTCCAGACTCGCTTACAGATAGTGTGTGATCTCACCTCTGAATATTAT 240
Db 1131 CCTACTAGATCGCTCCAGACTCGCTTACAGATAGTGTGTGATCTCACCTGAATATTAT 1190
QY 241 CCCAGTGAAGATGATGTTATGAAAGGAGAGACTCATTTAGTGCAGTGCCTGATTTCTGT 300
Db 1191 CCCAGTGAAGATGATGTTATGAAAGGAGAGACTCATTTAGTGCAGTGCCTGATTTCTGT 1250
QY 301 GACGCCACTAATATTCTATGAAGAAACACAGACATCTGGATAAATATCCATGACATCTT 360
Db 1251 GACGCCACTAATATTANCNATGAAGAAACACAGACATCTGGATAAATATCCATGACATCTT 1310
QY 361 TCATGTTTTTCCCAAGTCAAGAGGAAATTTGAGTTTATTTTGCCTCTGAATGCAA 420
Db 1311 TCATGTTTTTCCCAAGTCAAGAGGAAATTTGAGTTTATTTTGCCTCTGAATGCAA 1370
QY 421 AACAGGTTTCGTCATTTATACAAATATCATCTATTTTAAAGGAAAGCAATATAAAG 480
Db 1371 AACAGGTTTCGTCATTTATACAAATATCATCTATTTTAAAGGAAAGCAATATAAAG 1430
QY 481 ATCCAGTGTGGGCTGCTCTCCAGAGTATTTCAAGTGTCTCTATCAAGAGAGATAGC 540
Db 1431 ATCCAGTGTGGGCTGCTCTCCCAANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1490
QY 541 AATTACCAGTGTGATGGAAGTCTCTGGCCGCATGATCTAATATATCCAAAGTTGATGA 600
Db 1491 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1550
QY 601 AGTCAGAGGCTGATATATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACCTGTA 660
Db 1551 AGTCAGAGGCTGATATATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACCTGTA 1610
QY 661 CGTAGTCAGTTACGTAATCTCTGAGAGGTGACAGGCTGACTGACCGTGGCTACTCACA 720
Db 1611 CGTAGTCAGTTACGTAATCTCTGAGAGGTGACAGGTTGACTGACCGTGGCTACTCACA 1670
QY 721 TTCTTGCTGCTCAGTCAGCATCTGACTTCTTTTAAAGTAAAGTATAGTAACCAAGAA 780
Db 1671 TTCTTGCTGCTCAGTCAGCATCTGACTTCTTTTAAAGTAAAGTATAGTAACCAAGAA 1730
QY 781 TCCACACTGTGTGCTCCCTTTTCAAGCTATCAAGTCTCTGAAGATGACCCCACTTGCAAAAC 840
Db 1731 TCCACACTGTGTGCTCCCTTTTCAAGCTATCAAGTCTCTGAAGATGACCCCACTTGCAAAAC 1790
QY 841 AAAGGAATTTTGGGCGACCATTTTGGATTTCAGCAGGTCTCTTCTGACTATATCTCTCC 900
Db 1791 AAAGGAATTTTGGGCGACCATTTTGGATTTCAGCAGGTCTCTTCTGACTATATCTCTCC 1850
QY 901 AGAAATTTCTCTTTTGAAGTACTACTGATTTACATTTGATGGATGCTCTACAAGCC 960
Db 1851 AGAAATTTCTCTTTTGAAGTACTACTGATTTACATTTGATGGATGCTCTACAAGCC 1910
QY 961 TCATGATCTACAGCTGGAAGAAATATCTCTGCTGTGTTTCATATATGTTGGTGGTCTCA 1020
Db 1911 TCATGATCTACAGCTGGAAGAAATATCTCTGCTGTGTTTCATATATGTTGGTGGTCTCA 1970
QY 1021 ----- 1020
Db 1971 GGTGAGTTGGTGAATAAATCGGTTTAAAGGAGTCAAGTATTTCCGTTGAATACCTTAGC 2030
QY 1021 ----- 1020
Db 2031 CTCTCTAGGTTATGTTGTTAGTGTAGTATAGAACAGGGGATCTCTGTCCCGAGGGCTTAA 2090
```

QY	1021	-----	1020
Db	2091	ATTGTAAGGCGCCTTTAAATATATAAATGGTCAAAATAGAAATTCAGCATCAGTGGGAAGG	2150
QY	1021	-----	1020
Db	2151	ACTCCAATATCTAGCTTCTCGATATGATTTTCATTGACTTAGATCGTGGGCATCCACGG	2210
QY	1021	-----	1020
Db	2211	CTGTCTCTATGGAGGATACCTCTCCCTGATGGCAITTAATGCAGAGGTTCAGATATCTTCAG	2270
QY	1021	GGTTGCTATTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGA	1080
Db	2271	GGTTGCTATTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGA	2330
QY	1081	ACGTTATATGGGTCAACCTTGACCAGAAATGAACAGGCTATTACTTAGGATCTGTGGCCAT	1140
Db	2331	ACGTTATATGGGTCAACCTTGACCAGAAATGAACANNNNNNNNNNNNNNNNNNNNNN	2390
QY	1141	GCAAGCAGAAAGTTCCTCTGAACCAATCGTTTACTGCTCTTACTATGATGTTCTCTGGA	1200
Db	2391	NN	2450
QY	1201	TCAGATGTCATTTGACATACACAGTATATTAATGATGTTTCTGCTCTTACTGCTCTTCTGGA	1260
Db	2451	TGAGATGTCATTTGACATACACAGTATATTAATGATGTTTCTGCTCTTACTGCTCTTCTGGA	2510
QY	1261	GCCATATGATTTACAGATCTATCTCTCAGGAGACACAGCATAGAGTTCCTGAATCGGG	1320
Db	2511	ACCATATGATTTACAGATCTATCTCTCAGGAGACACAGCATAGAGTTCCTGAATCTGG	2570
QY	1321	AGAACTATTGAATGCATCTTTTGCACTACCTTCAAGAAACCTTGATGATCAGATATGC	1380
Db	2571	AGAACTATTGAATGCATCTTTTGCACTACCTTCAAGAAACCTTGATGATCAGATATGC	2630
QY	1381	TGCTCTAAAAGTGATATAA	1399
Db	2631	TGCTCTAAAAGTGATATAA	2649
RESULT 6			
BX372276/G			
LOCUS			
DEFINITION BX372276 Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED			
Homo sapiens cDNA clone CS0DL005YD02 3-PRIME, mRNA sequence.			
BX372276			
VERSION BX372276.1 GI:30448117			
KEYWORDS EST			
SOURCE Homo sapiens (human)			
ORGANISM Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.			
1 (bases 1 to 910)			
Li,W.B., Gruber,C., Jessee,J. and Polayes,D.			
Full-length cDNA libraries and normalization			
Unpublished (2001)			
Contact: Genoscope			
Genoscope - Centre National de Sequencage			
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE			
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr			
1st strand cDNA was primed with a NotI-Oligo(dT) primer. Five prime			
end enriched, double-strand cDNA was digested with Not I and cloned			
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library			
was normalized. Library was constructed by Life Technologies, a			
division of Invitrogen. This sequence belongs to sequence cluster			
7542.r			
For more information about this cluster, see			
http://www.genoscope.cns.fr/cdna?s=CS0BAI043ZH06_CS04080_1&c=7542.r			
. FEATURES			
Location/Qualifiers			

source		1. 910	
		/organism="Homo sapiens"	
		/mol_type="mRNA"	
		/db_xref="taxon:9606"	
		/clone="CS0DL005YD02"	
		/cell_type="B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"	
		/cell_line="RAMOS CELL LINE"	
		/clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"	
		/notes="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."	
ORIGIN		Query Match	
		Best Local Similarity 49.2%; Score 820.6; DB 5; Length 910;	
		Matches 856; Conservative 99.0%; Pred. No. 4.2e-201; Indels 3; Gaps 3;	
QY	1	AACAGGTACAGCAAACTCTAAAGTCACTTTTAAGATGTCAAGAAATAATGATTGATCTGA	60
Db	870	AACAGGTACAGCAAACTCTAAAGTCACTTTTAAGATGTCAAGAAATAATGATTGATCTGA	811
QY	61	AGGAGGATCATAGATGTCATAGATAAGAACTAATTCACCTTTTGAGATTCATTATGA	120
Db	810	AGGAGGATCATAGATGTCATAGATAAGAACTAATTCACCTTTTGAGATTCATTATGA	753
QY	121	AGGAGTTGAATATATTGCCAGAGCTGGATGGACTCTCTGAGGAAATAATGCTTGGTCCAT	180
Db	752	AGNAGTTGGATATATTGCCAGAGCTGGATGGACTCTCTGAGGAAATAATGCTTGGTCCAT	693
QY	181	CCTACTAGATCGCTCCACAGACTCGCTACAGATAGTTGTGATCTCACCTGAATATTATAT	240
Db	692	CCTACTAGATCGCTCCACAGACTCGCTACAGATAGTTGTGATCTCACCTGAATATTATAT	633
QY	241	CCCAGTAAAGATGATGTTATGGAAGGAGAGAGATCATTGAGTCAGTCGCTGATTCGT	300
Db	632	CNCAGTAAAGATGATGTTATGGAAGGAGAGAGATCATTGAGTCAGTCGCTGATTCGT	573
QY	301	GACGCCACTAATATTCTATGAAGAAACAAACAGACATCTGGATAAATATCCATGACATCTT	360
Db	572	GACGCCACTAATATTCTATGAAGAAACAAACAGACATCTGGATAAATATCCATGACATCTT	513
QY	361	TCATGTTTTTCCCAAGTCACGAGAGGAAATTCAGTTTATTTTGGCTCTGATGCAA	420
Db	512	TCATGTTTTTCCCAAGTCACGAGAGGAAATTCAGTTTATTTTGGCTCTGATGCAA	453
QY	421	AACAGGTTTCCGTCATTATACAAATTAATACATCTATTTAAAGGAAACCAATATAACG	480
Db	452	AACAGGTTTCCGTCATTATACAAATTAATACATCTATTTAAAGGAAACCAATATAACG	393
QY	481	ATCCAGTGGTGGCTGCCTGCTCCAGTCATTTCAAGTGTCTATCAAGGAGGATAGC	540
Db	392	ATCCAGTGGTGGCTGCCTGCTCCAGTCATTTCAAGTGTCTATCAAGGAGGATAGC	333
QY	541	AATTACAGTGGTGAATGGGAGTTCTTCCGCGCATGGATCTAATATCCAAAGTTGATGA	600
Db	332	AATTACAGTGGTGAATGGGAGTTCTTCCGCGCATGGATCTAATATCCAAAGTTGATGA	273
QY	601	AGTCAGAGCTCGGTATATTTTGAAGGCAACAAAGACTCCCTTTAGAGCATCACCTGTA	660
Db	272	AGTCAGAGCTCGGTATATTTTGAAGGCAACAAAGACTCCCTTTAGAGCATCACCTGTA	213
QY	661	CGTAGTCAGTTACGTAATTCCTGGAGAGGTGACAAAGGCTGACTGACCGTGGCTACTACA	720
Db	212	CGTAGTCAGTTACGTAATTCCTGGAGAGGTGACAAAGGCTGACTGACCGTGGCTACTACA	153
QY	721	TTCTTGTGTCATCAGTCAGCACTGTGACTCTTTTATAAGTAAGTATAGTACCAAGAGAA	780
Db	152	TTCTTGTGTCATCAGTCAGCACTGTGACTCTTTTATAAGTAAGTATAGTACCAAGAGAA	93
QY	781	TCCACACTGTGTCTCCCTTTTACAAGCTATCAAGTCTCTGAAGATGACCCCACTTGCAAAAC	840

```

Db      92  TCCACACTGTGTCCCTTTACAAGCTATCAAGTCTGAAGATGACCCAAAC-TGCAAAAC 34
Qy      841 AAAGGAATTTTGGGCCACCAATTTTG 865
Db      33  AAAGGAATTTTGGGCCACCAATTTGG 9

RESULT 7
DN517062/c
LOCUS      DN517062      902 bp      mRNA      linear      EST 10-MAR-2005
DEFINITION 1257572 MARC 7BOV Bos taurus cDNA 3', mRNA sequence.
ACCESSION  DN517062
VERSION     DN517062.1  GI:60727252
KEYWORDS    EST.
SOURCE      Bos taurus (cow)
ORGANISM    Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
            Pecora; Bovidae; Bovinae; Bos.
REFERENCE   1 (bases 1 to 902)
AUTHORS     Smith,T.P.L., Roberts,A.J., Echternkamp,S.E., Chitko-McKown,C.G.,
            Wray,J.E. and Keefe,J.W.
TITLE       A second set of bovine ESTs from pooled-tissue normalized libraries
JOURNAL     Unpublished (2003)
COMMENT     Contact: Smith TPL
            USDA, ARS, US Meat Animal Research Center
            PO Box 166, Clay Center, NE 68933-0166, USA
            Tel: 402 762 4366
            Fax: 402 762 4390
            Email: smith@email.marc.usda.gov
            Single pass sequencing. Bases called with phred v0.020425.c and
            trimmed with the aid of the trim_alt option. Vector identified with
            cross_match v0.990329.
            Seq primer: TAGAGGCACAGTCGAGG.
FEATURES             source
            1..902      Location/Qualifiers
                        /organism="Bos taurus"
                        /mol_type="mRNA"
                        /db_xref="taxon:9913"
                        /tissue_type="pooled"
                        /lab_host="DH10B"
                        /clone_lib="MARC 7BOV"
                        /notes="Vector: pcDNA3.1; Site 1: EcoRI; Site 2: NotI;
                        Library made with RNA pooled from multiple tissues
                        including ovary, hindbrain, uterus, and day-30 whole
                        embryos."
ORIGIN
Query Match      48.2%; Score 804.4; DB 8; Length 902;
Best Local Similarity 93.2%; Pred. No. 6.e-197;
Matches 841; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

Qy      4  AGGTACAGCAATCTTAAGTCACTTTTAAGATGTCAGAAATATGATGATGCTGAAGG 63
Db      902 AGGCACAGCAATCAAAAAGTCACTTTTAAGATGTCAGAAATATGATGATGCTGAAGG 843

Qy      64  AAGGATCATAGATGTCATAGATAAGGAACCTAATTCACCTTTTGAGATTCCTATTGTAAGG 123
Db      842 AAGGATTAATAGATGTCATAGATAAGGAACCTAATTCAGCCTTTTGAGATTCCTATTGTAAGG 783

Qy      124 AGTTGAATATATTGCCAGAGCTGGATGGACTCTCTGAGGGAAAATATGCTTGCTCCATCCT 183
Db      782 AGTTGAATATATTGCCAGAGCTGGATGGACTCCAGGGGAAAATATGCTTGCTGGCCATCCT 723

Qy      184 ACTAGATCGCTCCAGACTCGCCTACAGATAGTGTGTGATCTCACCTGAATATTATTATCCC 243
Db      722 ACTAGATCGCTCCAGACTCGCCTACAGATAGTGTGTGATCTCACCTGAATATTATTATCCC 663

Qy      244 AGTAGAGATGATGTTATGAAAGGACAGACATCATTTAGTCAGTGCCTGATTCGTGAC 303
Db      662 AGTAGAGATGATGTCATGAAAGACAGAGATTTCATCGAGTCAGTGCCTGATTCGTGAC 603

```

```

Qy      304 GCCACTAATATTCTATGAAGAAACAAACAGACATCTGGATAAAATATCCATGACATCTTTCA 363
Db      502 ACCACTGATTAATCTATGAAGAAACGACAGACATCTGGATAAAATATCCATGACATCTTTCA 543
Qy      364 TGTGTTTCCCAAAAGTCAAGAGGAAATGAGTTTATTTTGGCTCTGATCCTGATCCTGATCCT 423
Db      542 TGTGTTTCCCAAAAGTCAAGAGATGAATTTGAATTTATTTTGGCTCTGATCCTGATCCTGATCCT 483
Qy      424 AGGTTTCCGCTATTATACAAAATACATCTATTTTAAAGGAAAGCAATATAAACGATC 483
Db      482 AGGTTTCCGCTATTATATANAATCACATCCATTTTAAAGGAGAGCAAGTATAAGCCATC 423
Qy      484 CAGTGTGGGCTGCTGCTCCCAAGTCAATTTCAAGTGTCTCTATCAAAAGAGGAGATAGCAAT 543
Db      422 CAGTGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 363
Qy      544 TACCAGTGTGTGAATGGAAAGTCTTTGGCGGCGATCGATCTAATATPCCAAAGTTGATGAAGT 603
Db      362 TACCAGTGTGTGAATGGAAAGTCTTTGGCGGCGATCGATCTAATATPCCAAAGTTGATGAAGT 303
Qy      604 CAGAAGCTGGTATATATTTTGAAGGACCAAAAGACTCCCTTTAGAGCATCACCTGTAGCT 663
Db      302 CAGAAGCTGGTATATATTTTGAAGGACCAAAAGACTCTCTTTAGAACATCACCTGTATGT 243
Qy      664 AGTCAGTTAGTAAATCCTGGAGAGGTGACAGGCTGACCTGACCTGCTGCTGCTGCTGCTGCTGCT 723
Db      242 GGTCAAGTTATGTAATCCTGGAGAGGTGACAGGCTGACCTGACCTGCTGCTGCTGCTGCTGCTGCT 183
Qy      724 TTGCTGCATCAGTCAGCACTGTGACTTCTTTAAGTAAGTATAGTACCAAGAAATCC 783
Db      182 TTGCTGCATCAGTCAGCGCATGCTGACTTCTTTAAGTAAGTATAGTACCAAGAAAGGCCCC 123
Qy      784 ACACGTGTGCTCCCTTTTACAAGCTATCAAGTCTCTCAAGATGACCCCACTTCGCAAAACAAA 843
Db      122 ACACGTGTGCTCCCTTTTACAAGCTCTCAAGTCTCTCAAGCAGCACCACCTGCAAGACAAA 63
Qy      844 GGAATTTTGGGCCACCACTTTTGGATTCAGCAGGTCTCTTCTGCTGATCTATCTCTCCAGA 903
Db      62 AGAATCTGCGGCCACCGCTCTTTGGATTTCAGCAGGTCTCTTCTGCTGATCTATACCCCTCCAAA 3
Qy      904 AA 905
Db      2 AA 1

RESULT 8
DN519771
LOCUS      DN519771      857 bp      mRNA      linear      EST 11-MAR-2005
DEFINITION 1262448 MARC 7BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION  DN519771
VERSION     DN519771.1  GI:60958943
KEYWORDS    EST.
SOURCE      Bos taurus (cow)
ORGANISM    Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
            Pecora; Bovidae; Bovinae; Bos.
REFERENCE   1 (bases 1 to 857)
AUTHORS     Smith,T.P.L., Roberts,A.J., Echternkamp,S.E., Chitko-McKown,C.G.,
            Wray,J.E. and Keefe,J.W.
TITLE       A second set of bovine ESTs from pooled-tissue normalized libraries
JOURNAL     Unpublished (2003)
COMMENT     Contact: Smith TPL
            USDA, ARS, US Meat Animal Research Center
            PO Box 166, Clay Center, NE 68933-0166, USA
            Tel: 402 762 4366
            Fax: 402 762 4390
            Email: smith@email.marc.usda.gov
            Single pass sequencing. Bases called with phred v0.020425.c and
            trimmed with the aid of the trim_alt option. Vector identified with
            cross_match v0.990329.
            Seq primer: GTAATACGACTCCTATAGG.

```

```
FEATURES
  source      Location/Qualifiers
1..857
  /organism="Bos taurus"
  /db_xref="taxon:9913"
  /tissue_type="pooled"
  /lab_host="DH10B"
  /clone_lib="MARC 780V"
  /notes="Vector: pCDNA3.1; Site 1: EcoRI; Site 2: NotI;
  Library made with RNA pooled from multiple tissues
  including ovary, hindbrain, uterus, and day-30 whole
  embryos."

ORIGIN
Query Match      43.8%; Score 730.6; DB 8; Length 857;
Best Local Similarity 93.9%; Pred. No. 8.5e-178;
Matches 760; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

Qy 1 AACAGGTACAGCAAAATCCATAAGTCACTTTTAAGATGTGAGAAATATGATGATGCTGA 60
Db 49 AACAGGCACAGCAAAATCCAAAAGTCACTTTAAAGATGTGAGAAATATGATGATGCTGA 108
Qy 61 AGGAAGGATCATAGATGTCTAGATAAGGAACATAATTCACCTTTTTCAGATTCTATTGTA 120
Db 109 AGGAAGGATTTATAGATGTCTAGATAAGGAACATAATTCAGCCTTTTTCAGATTCTATTGTA 168
Qy 121 AGGAGTTGAATATATTGTCAGAGCTGGATGGAATCTCTGAGGGAATAATGCTTGGCCAT 180
Db 169 AGGAGTTGAATATATTGTCAGAGCTGGATGGAATCTCTGAGGGAATAATGCTTGGCCAT 228
Qy 181 CCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAATATTAT 240
Db 229 CCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAATATTAT 288
Qy 241 CCCAGTAGAGATGATGTTATCGGAAGCGAGAGACTCATTTGAGTCAGTGCCTGATCTGT 300
Db 289 CCCAGTAGAGATGATGCTCATGGAAGACAGAGATTCTCATGTCAGTGCCTGATCTGT 348
Qy 301 GAGCCACTAATTAATCTATGAGAAACACAGACATCTGGAATAATATCCATGACATCTT 360
Db 349 GACACCACTGATTAATCTATGAGAAACACAGACATCTGGAATAATATCCATGACATCTT 408
Qy 361 TCATGTTTTTCCCAAGTCACGAAGAGAAATTTGAGTTTATTTTTTGCCTCTGAATGCAG 420
Db 409 TCATGTTTTTCCCAAGTCATGAAGATGAAATTTGAAATTTATTTTTTGCCTCTGAATGCAG 468
Qy 421 AACAGTTTCCGTCAATTTATACAAATTTACATCTATTTTAAAGGAAGCAAAATATAAAGC 480
Db 469 AACAGTTTCCGTCAATTTATATAAATACATCTATTTTAAAGGAAGCAAGTATAAGCG 528
Qy 481 ATCCAGTGTGGGCTGCCTGCTCCAAAGTGATTTTCAAGTGTCTTCAAGAGGAGATAGC 540
Db 529 ATCCAGTGTGGGCTGCCTGCTCCCAAGTGATTTTCAAGTGTCTTCAAGAGGAAATAGC 588
Qy 541 AATTACAGTGTGAATGGGAAGTTCTTGGCCGCATGGATCTAATATCCAGTTGATCA 600
Db 589 AATTACAGTGTGAATGGGAAGTTCTTGGCCGCATGGATCTAATATCCAGTTGATCA 648
Qy 601 AGTCAGAGGCTGGTATATTTTGAAGGCACCAAGACTCCCTTTAGAGCATCACTGTA 660
Db 649 AGTCAGAAAGCTGGTATATTTTGAAGGCACCAAGACTCTCTTTAGACATCACTGTA 708
Qy 661 CGTAGTCAGTACGTAAATCTCTGAGAGGTGACAGAGCTGACGCGGTGCTACTACA 720
Db 709 TGTGTCAGTATATGTAATCTCTGAGAGGTGACAGAGCTGACGCGGTGCTACTACA 768
Qy 721 TTCTTGTGTCATCAGTCAGCACTGTGACTTCTTTTAAAGTAAGTAAACAGAGAA 780
Db 769 CTCCTGTCATCAGCCGGCATTTGTGACTTCTTTTAAAGTAAGTAAACAGAGAA 828
Qy 781 TCCACACTGTGTGCTCCCTTACAGCTAT 809
Db 829 CCCACACTGTGTGCTCCCTTACAGCTGT 857
```

```
RESULT 9
B1223892
LOCUS      B1223892
DEFINITION      602941035f1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:5104113 5',
mRNA sequence.
ACCESSION      B1223892
VERSION        B1223892
KEYWORDS       EST.
SOURCE         Homo sapiens (human)
ORGANISM       Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
1 (bases 1 to 853)
NIH-MGC http://mgs.nci.nih.gov/.
NATIONAL INSTITUTE OF HEALTH, MAMMALIAN GENE COLLECTION (MGC)
UNPUBLISHED (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: Incyte Genomics, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LHAM1251 row: e column: 10
High quality sequence stop: 718.
Location/Qualifiers
1..853
  /organism="Homo sapiens"
  /mol_type="mRNA"
  /db_xref="taxon:9606"
  /clone="IMAGE:5104113"
  /tissue_type="cervical carcinoma cell line"
  /lab_host="DH10B"
  /clone_lib="NIH MGC_12"
  /notes="Organ: cervix; Vector: pCMV-SPORT6; Site 1: NotI;
  Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
  Average insert size 1.4 kb. Library prepared by Life
  Technologies."
```

```
ORIGIN
Query Match      41.2%; Score 68814; DB 2; Length 853;
Best Local Similarity 99.7%; Pred. No. 7.3e-167;
Matches 700; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy 323 GAAACCAACAGACATCTGGATAAATATCCATGACATCTTTCA-TGTTTTTCCCAAGTCA 381
Db 1 GAAACCAACAGACATCTGGATAAATATCCATGACATCTTTCAAGTGTGTTTCCCAAGTCA 60
Qy 382 CGAAGAGGAAATTTGAGTTTATTTTGGCTCTGAATGCAAAACAGGTTTCGTCATTTATA 441
Db 61 CGAAGAGGAAATTTGAGTTTATTTTGGCTCTGAATGCAAAACAGGTTTCGTCATTTATA 120
Qy 442 CAAATTTACATCTATTTTAAAGGAAGCAATATAACGATCCAGTGGTGGCTGCTGC 501
Db 121 CAAATTTACATCTATTTTAAAGGAAGCAATATAACGATCCAGTGGTGGCTGCTGC 180
Qy 502 TCCAGTGAATTTCAAGTGTCTTATCAAAAGAGGAGATAGCAATTAACAGTGGTGAATGGGA 561
Db 181 TCCAGTGAATTTCAAGTGTCTTATCAAAAGAGGAGATAGCAATTAACAGTGGTGAATGGGA 240
Qy 562 AGTTCTTGGCCGGCATGGATCTTAATATCCAAAGTTGATGAAGTCAGAAAGCTGGTATATTT 621
Db 241 AGTTCTTGGCCGGCATGGATCTTAATATCCAAAGTTGATGAAGTCAGAAAGCTGGTATATTT 300
Qy 622 TGAAGSCACCAAGAACTCCCTTTTAGAGCATCACCTGTACGTAGTCAAGTTACGTAATCC 681
Db 301 TGAAGSCACCAAGAACTCCCTTTTAGAGCATCACCTGTACGTAGTCAAGTTACGTAATCC 360
```

Qy	682	TGAGAGGTGACAAAGCTGACCTGACCGTGGCTACTCACTCTTGGTGCATCAGTACGCA	741
Db	361	TGAGAGGTGACAAAGCTGACCTGACCGTGGCTACTCACTCTTGGTGCATCAGTACGCA	420
Qy	742	CTGTGACTTCTTTATTAAGTAAGTATAGTAACGAGAGAAATCCACACTGTGTGCTCCCTTTA	801
Db	421	CTGTGACTTCTTTATTAAGTAAGTATAGTAACGAGAGAAATCCACACTGTGTGCTCCCTTTA	480
Qy	802	CAAGCTATCAAGTCTCTGAAGATGACCCAACTTGCAGAAACAAAGGAATTTTGGGCGCCAT	861
Db	481	CAAGCTATCAAGTCTCTGAAGATGACCCAACTTGCAGAAACAAAGGAATTTTGGGCGCCAT	540
Qy	862	TTTGGATTACGAGGTCCTCTTCCTGACTATATCTCCTCAGAAATTTCTTTTGAAG	921
Db	541	TGTGGATTACGAGGTCCTCTTCCTGACTATATCTCCTCAGAAATTTCTTTTGAAG	600
Qy	922	TACTACTGGATTACATTTGATGGGATGCTCTACAGGCTCATGATCTACAGCTGGAAA	981
Db	601	TACTACTGGATTACATTTGATGGGATGCTCTACAGGCTCATGATCTACAGCTGGAAA	660
Qy	982	GAATATCTCTACTGTGCTGTTTCATATATGTTGGTCTCTCAGGT	1023
Db	661	GAATATCTCTACTGTGCTGTTTCATATATGTTGGTCTCTCAGGT	702
RESULT 10			
LOCUS	AY411617	2634 bp	DNA linear
DEFINITION	Mus musculus DPP8 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.		
ACCESSION	AY411617		
VERSION	AY411617.1		
KEYWORDS	GSS.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
AUTHORS	Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarawal, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.		
TITLE	Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios		
JOURNAL	Science 302 (5652), 1960-1963 (2003)		
PUBMED	14671302		
REFERENCE	2 (bases 1 to 2634)		
AUTHORS	Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarawal, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.		
TITLE	Direct Submission		
JOURNAL	Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA		
COMMENT	This sequence was made by sequencing genomic exons and ordering them based on alignment		
FEATURES	Location/Qualifiers		
source	1. .2634		
	/organism="Mus musculus"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:10090"		
gene	<1..2634		
	/gene="DPP8"		
	/locus_tag="HCM4283"		
ORIGIN			
Query Match	40.5%	Score 675.6;	DB 10; Length 2634;
Best Local Similarity	72.1%	Pred. No. 2e-163;	
Matches 738;	Conservative 0;	Mismatches 285;	Indels 0; Gaps 0;
Qy	1	AACAGGTACAGCAATCTTAAAGTCACTTTTAAGATGTGCAAAATTAATGATTGCTGA	60

Db	954	AACAGGCACAGCAAAACCCAAAGGTCACTTTCAAGATGTCGGAGATTGTTGATGCTGC	1013
Qy	61	AGGAAGGATCATAGATGTATAGATAAGCACTAAATTCACCTTTTGAGATTTCTATTGA	120
Db	1014	AGGAGGCGNN	1073
Qy	121	AGGAGTTGAATATATTGCGAGCTGGATGGACTCTCTGAGGAAATATGCTTGGTCCAT	180
Db	1074	NN	1133
Qy	181	CCTACTAGATCGCTCCAGAGCTCGCTACAGATAGTGTGTGATCTCACCTGTAATTTAT	240
Db	1134	CCTACTAGACCGTTCCAGACTCACTGCGAGATGTTCTGATCTCCCTGAGTATTCTAT	1193
Qy	241	CCCAGTGAAGATGATGTTATGGAAGGAGAGACTCATTTGAGTCACTGCTGATTTCTGT	300
Db	1194	CCCAGTGAAGATGATGCTCATGACAGAGACTTATAGAGTCACTGCTGACTCTGT	1253
Qy	301	GACGCCATTAATTTCTATGAAGAAACACAGACATCTGGATAAATATCCATGACATCTT	360
Db	1254	GACACCACTGATCATCTATGAAGAAACACAGACATCTGGATAAATATCCACGATATTT	1313
Qy	361	TCATGTTTTTCCCAAGTCAACGAGAGAAATGAGTTTATTTTGGCTCTGAATGCAA	420
Db	1314	TCATGTTTTTCCCTCAAACTCTGAAAGATGAATTTGAGTTTATTTTGGCTCTGAATGCAA	1373
Qy	421	AACAGGTTTCCGTCATTATACAAATTTACATCTATTTTAAAGGAAAGCAATATAAAG	480
Db	1374	AACAGGTTTTCGTCATCTGTATAAATCATATCCATTTTAAAGGAGCAAAATATAAAG	1433
Qy	481	ATCAGTGTGGTGGCTGCTCTCCAAAGTGAATTTCAAGTGTCTATCAAGAGGAGATAGC	540
Db	1434	GTCCAGTGTGGTGGTCTCTCCCAANNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	1493
Qy	541	AATTACCAGTGTGAATGGGAAGTTCTTGTGCGGCATGGATCTAATATCCAAGTTGATGA	600
Db	1494	NN	1553
Qy	601	AGTCAGAGGCTGGTATATTTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTA	660
Db	1554	AGCCAGAAAGCTGGTGTACTTTGAGGCACCAAGACTCTCTTTGGAGCATCACCTGTA	1613
Qy	661	CGTGTGAGTTACGTAATCTCTGAGAGGTGACAGAGCTGACTGACCGTGGTACTCACA	720
Db	1614	CGTCACCAAGTTATGCAAAACCTGAGAGAGTGTGTGAGGCTGACTGACCGTGGTACTCACA	1673
Qy	721	TTCTTGTGCTGATCAGTCAGCATCTGACTTCTTTTATAAGTAAAGTATAGTAACCAAGAA	780
Db	1674	CTCTGTGCTGCTCAGCGGCAATTTGACTTCTTTTATAAGTAAAGTACAGCAACCAAGAA	1733
Qy	781	TCCACACTGTGTGTCCTTTTACAAGCTATCAAGTCTCTGAAGATGACCCCAACTTGCAGAA	840
Db	1734	TCCACACTGTGTGTCCTTTTACAAGCTATCAAGTCTCTGAAGATGACCCCAACTTGCAGAA	1793
Qy	841	AAAGGAAATTTTGGGCCACCAATTTTGGATTACAGAGTCTCTTCTGACTATACCTCTCC	900
Db	1794	AAAGGAAATTTTGGGCCACCAATTTTGGATTACAGAGTCTCTTCTGACTATACCTCTCC	1853
Qy	901	AGAAATTTTCTCTTTTGAAGTACTACTGATTTTACATTTGATGGATGCTCTACAAGCC	960
Db	1854	AGAAATTTTCTCTTTTGAAGTACTACTGATTTTACATTTGATGGATGCTCTACAAGCC	1913
Qy	961	TCATGATCTACAGGCTGGAAGAAATATCTACTGTGCTGTTTCATATATGGTGGTCTCA	1020
Db	1914	TCATGATCTACAGGCTGGAAGAAATATCTACTGTGCTGTTTCATATATGGTGGTCTCA	1973
Qy	1021	GCT.1023	
Db	1974	GCT 1976	
RESULT 11			
CN647232			

LOCUS CN647232 1013 bp mRNA linear EST 13-MAY-2004
DEFINITION ILLUMIGEN MQC 28042 Katze MBR Macaca mulatta cDNA clone IB1UW:7816
5' similar to Bases 1 to 1009 highly similar to human DPP8
(Hs.439202), mRNA sequence.
ACCESSION CN647232
VERSION CN647232.1 GI:47160675
KEYWORDS EST
SOURCE Macaca mulatta (rhesus monkey)
ORGANISM Macaca mulatta
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Cercopithecoidea; Cercopithecinae; Macaca.
REFERENCE 1 (Bases 1 to 1013)
AUTHORS Magness, C.L., Fellin, P.C., Thomas, M.J., Korth, M.J., Agy, M.B.,
Proli, S.C., Fitzgibbon, M., Scherer, C.A., Miner, D.G., Katze, M.G. and
Iadonato, S.P.
TITLE Analysis of the Macaca mulatta transcriptome and the sequence
divergence between Macaca and human
JOURNAL Genome Biol. 6 (7), R60 (2005)
PUBMED 15998449
COMMENT Contact: C. Magness
Illumigen Biosciences Inc.
2203 Airport Way S, Suite 450, Seattle, WA 98134, USA
Tel: 2063780400
Fax: 2063780408
Email: cmagness@illumigen.com
PCR Primers
Sequenced on 2004.03.20. 793 Q20 bases.
FORWARD: CCTCACTAAAGGGAACAAA
BACKWARD: CACTATAGGCGAATTGGTA
Insert Length: 1013 Std Error: 0.00
Plate: CL000179 Row: E Column: 08
Seq primer: CCTCACTAAAGGGAACAAA
POLYA=No Location/Qualifiers
1..1013
/organism="Macaca mulatta"
/mol_type="mRNA"
/strain="Indian"
/db_xref="taxon:9544"
/clone="IB1UW:7816"
/sex="female"
/dev_stage="adult"
/lab_host="E. coli SOLR"
/clone_lib="Katze_MBR"
/notes="Organ: brain; Vector: Uni-ZAP XR; Site 1: EcoR I;
Site 2: Xho I; Created from Stratagene ZAP-CDNA Synthesis
kit (Catalog #200400) and ZAP-CDNA Gigapack III Gold
Cloning Kit (Catalog #200450)"

ORIGIN
Query Match 40.3%; Score 672.6; DB 7; Length 1013;
Best Local Similarity 98.7%; Pred. No. 9.4e-163;
Matches 678; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 337 CTGGATAAATATCCATGACATCTTTTCATGTTTTCCTCCCAAGTCACGAAGGAAATTGA 396
DB 1 CTGGATAAATATCCATGACATCTTTTCATGTTTTCCTCCCAAGTCACGAAGGAAATTGA 60
QY 397 GTTATTTTGGCTCTGAATGCAAAACAGGTTTCGGTCATTATACAAAATTACATCTAT 456
DB 61 GTTATTTTGGCTCTGAATGCAAAACAGGTTTCGGTCATTATACAAAATTACATCTAT 120
QY 457 TTATAGGAAGCAATATTAACGATCCAGTGGTGGCTGCTGCTCCCAAGTCAATTCAA 516
DB 121 TTATAGGAAGCAATATTAACGATCCAGTGGTGGCTGCTGCTCCCAAGTCAATTCAA 180
QY 517 GTGTCTATCAAGAGGAGATAGCAATTTACAGTGGTGAATGGGAAGTTCTTGGCCGGCA 576
DB 181 GTGTCTATCAAGAGGAGATAGCAATTTACAGTGGTGAATGGGAAGTTCTTGGCCGGCA 240
QY 577 TGATCTTAATATCCAAAGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGGCCACCAAGA 636

DB 241 TGGATCTAATATCAAGTTGATGAGTCAGAAGCTGGTATATATTTTGAAGGCCACCAAGA 300
QY 637 CTCCTCTTTAGAGCATCACCTGTACGTAGTACGTAAATCTCTGGAGAGGTGACAAAG 696
DB 301 CTCCTCTTTAGAGCATCATCTGTACGTAGTACGTAAATCTCTGGAGAGGTGACAAAG 360
QY 697 GCTGACTGAACCGTGGCTACTCAATTTCTGCTGATCAGTCACTGAGACTTCTTTAT 756
DB 361 GCTGACTGAACCGTGGCTACTCAATTTCTGCTGATCAGTCACTGAGACTTCTTTAT 420
QY 757 AAGTAAGTATAGTAACCAAGCAAGAAATCCACACTGTGTGTCCTTTTACAAGCTATCAAGTCC 816
DB 421 AAGTAAGTATAGTAACCAAGCAAGAAATCCACACTGTGTGTCCTTTTACAAGCTATCAAGTTC 480
QY 817 TGAAGATGACCCCAACTTGCACAAACAAAGAAATTTTGGGCCACCATTTTGGATTACAGCAGG 876
DB 481 TGAAGATGATCTAACTTGCACAAACAAAGAAATTTTGGGCCACCATTTTGGATTACAGCAGG 540
QY 877 TCCTCTTCTGACTATATCTCTCCAGAAATTTTCTTTTGAAGTACTACTTGGATTAC 936
DB 541 TCCTCTTCTGACTATATACCCCTCCAGAAATTTTCTTTTGAAGTACTACTTGGATTAC 600
QY 937 ATTCTATGGGATGCTCTTACAAAGCTCTCATCATCATCATCATCATCATCATCATCATCTGT 996
DB 601 ATTCTATGGGATGCTATACAAAGCTCTCATCATCATCATCATCATCATCATCATCATCTGT 660
QY 997 GCTGTTTCATATATGGTGGTCTCAGGT 1023
DB 661 GCTGTTTCATATATGGTGGTCTCAGGT 687

RESULT 12
LOCUS CN646612 1027 bp mRNA linear EST 13-MAY-2004
DEFINITION ILLUMIGEN MQC 26659 Katze MBR Macaca mulatta cDNA clone IB1UW:8421
5' similar to Bases 1 to 1025 highly similar to human DPP8
(Hs.439202), mRNA sequence.
ACCESSION CN646612
VERSION CN646612.1 GI:47160055
KEYWORDS EST
SOURCE Macaca mulatta (rhesus monkey)
ORGANISM Macaca mulatta
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Cercopithecoidea; Cercopithecinae; Macaca.
REFERENCE 1 (Bases 1 to 1027)
AUTHORS Magness, C.L., Fellin, P.C., Thomas, M.J., Korth, M.J., Agy, M.B.,
Proli, S.C., Fitzgibbon, M., Scherer, C.A., Miner, D.G., Katze, M.G. and
Iadonato, S.P.
TITLE Analysis of the Macaca mulatta transcriptome and the sequence
divergence between Macaca and human
JOURNAL Genome Biol. 6 (7), R60 (2005)
PUBMED 15998449
COMMENT Contact: C. Magness
Illumigen Biosciences Inc.
2203 Airport Way S, Suite 450, Seattle, WA 98134, USA
Tel: 2063780400
Fax: 2063780408
Email: cmagness@illumigen.com
PCR Primers
Sequenced on 2004.03.10. 784 Q20 bases.
FORWARD: CCTCACTAAAGGGAACAAA
BACKWARD: CACTATAGGCGAATTGGTA
Insert Length: 1027 Std Error: 0.00
Plate: CL000146 Row: D Column: 07
Seq primer: CCTCACTAAAGGGAACAAA
POLYA=No Location/Qualifiers
1..1027
/organism="Macaca mulatta"
/mol_type="mRNA"
/strain="Indian"
/db_xref="taxon:9544"

```
/clone="IBIUM:8421"
/sex="female"
/dev_stage="adult"
/lab_host="E. coli SOLR"
/clone_lib="Katze_MBR"
/notes="Organ: brain; Vector: Uni-ZAP XR; Site 1: Ecor I;
Site 2: Xho I; Created from Stratagene ZAP-cDNA Synthesis
kit (Catalog #200400) and ZAP-cDNA Gigapack III Gold
Cloning Kit (Catalog #200450)"

ORIGIN
Query Match 40.3%; Score 672.6; DB 7; Length 1027;
Best Local Similarity 98.7%; Pred. No. 9.4e-163;
Matches 678; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 337 CTGGATAAATATCCATGACATCTTTCATGTTTTTCCCAAGTCCAGAGGAGAAATTGA 396
DB 1 CTGGATAAATATCCATGACATCTTTCATGTTTTTCCCAAGTCCAGAGGAGAAATTGA 60
QY 397 GTTATTATTTTGGCTCTGAATCAAAACAGGTTTCGGTCATTTATACAAAATTACATCTAT 456
DB 61 GTTATTATTTTGGCTCTGAATCAAAACAGGTTTCGGTCATTTATACAAAATTACATCTAT 120
QY 457 TTTAAAGGAAGCAAAATATAACGATCCAGTGGTGGCTGCTCTCCAAAGTATTTCAA 516
DB 121 TTTAAAGGAAGCAAAATATAACGATCCAGTGGTGGCTGCTCTCCAAAGTATTTCAA 180
QY 517 GTGTCCTATCAAGAGGAGATAGCAATTACAGTGGTGAATGGGAAGTCTTGGCCGGCA 576
DB 181 GTGTCCTATCAAGAGGAGATAGCAATTACAGTGGTGAATGGGAAGTCTTGGCCGGCA 240
QY 577 TGGATCTAATATCCAGTGTAGCAAGTCCAGAGGCTGTATATTTGAAGCCACCAAGA 636
DB 241 TGGATCTAATATCCAGTGTAGCAAGTCCAGAGGCTGTATATTTGAAGCCACCAAGA 300
QY 637 CTCCTCTTTAGAGCATCACCTGTACGTAGTACGTTACGTAATCTCGAGAGGTGACAAG 696
DB 301 CTCCTCTTTAGAGCATCATCTGTACGTAGTACGTTACGTAATCTCGAGAGGTGACAAG 360
QY 697 GCTGACTGACCGTGGCTACTCATTCTTGTGCAATCAGTCAGCAGTGTGACTTCTTTAT 756
DB 361 GCTGACTGACCGTGGCTATTTCATTCTTGTGCAATCAGTCAGCAGTGTGACTTCTTTAT 420
QY 757 AAGTAAGTATAGTAACAGAGAAATCCACACTGTGTGTCCTTTTACAAGCTATCAAGTCC 816
DB 421 AAGTAAGTATAGTAACAGAGAAATCCACACTGTGTGTCCTTTTACAAGCTATCAAGTTC 480
QY 817 TGAAGATGCCCAACTTGCAGAAAGGAATTTTGGGCCACCATTTTGGATTCCAGCAGG 876
DB 481 TGAAGATGATCTAACTTGCAGAAACAAAGGAATTTTGGGCCACCATTTTGGATTCCAGCAGG 540
QY 877 TCCTCTCTCTGACTACTACTCTCCAGAAATTTCTCTTTGAAAGTACTACTGATTTAC 936
DB 541 TCCTCTCTCTGACTACTACTCCCTCCAGAAATTTCTCTTTGAAAGTACTACTGATTTAC 600
QY 937 ATTGTATGGATGCTCTTACAAGCCTCATGATCTACAGCCTGGAAGAAATATCTACTGT 996
DB 601 ATTGTATGGATGCTATACAGGCTCATGATCTACAGCCTGGAAGAAATATCTCAACTGT 660
QY 997 GCTGTTTCATATGGTGGTCTCTCAGGT 1023
DB 661 GCTGTTTCATATGGTGGTCTCTCAGGT 687

RESULT 13
BG259714
LOCUS
DEFINITION 602380072F1 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:4511080 5',
mRNA sequence.
ACCESSION BG259714
VERSION BG259714.1 GI:12769530
KEYWORDS EST.
SOURCE Homo sapiens (human)
```

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
1 (bases 1 to 993)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM10393 row: 0 column: 17
High quality sequence start: 6
High quality sequence stop: 656.
Location/Qualifiers
1..993
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4511080"
/tissue_type="embryonal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_92"
/notes="Organ: testis; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dr primed.
Average insert size 2.5 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

ORIGIN
Query Match 39.5%; Score 660; DB 2; Length 993;
Best Local Similarity 100.0%; Pred. No. 1.7e-159;
Matches 660; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 277 CATTGAGTCAGTCGCTGATTTCTGTGAGCGCACATAATTATCTATGAAGAAACACAGACAT 336
DB 1 CATTGAGTCAGTCGCTGATTTCTGTGAGCGCACATAATTATCTATGAAGAAACACAGACAT 60
QY 337 CTGATATAATCCATGACATCTTTCATGTTTTTCCCAAGTCCAGAGGAAATTGA 396
DB 61 CTGATATAATCCATGACATCTTTCATGTTTTTCCCAAGTCCAGAGGAAATTGA 120
QY 397 GTTATTATTTTGGCTCTGAATCAAAACAGGTTTCCGTCAATTTATACAAAATTACATCTAT 456
DB 121 GTTATTATTTTGGCTCTGAATCAAAACAGGTTTCCGTCAATTTATACAAAATTACATCTAT 180
QY 457 TTTAAAGGAAGCAAAATATAAAGATCCAGTGGTGGCTGCTCTCCAAAGTATTTCAA 516
DB 181 TTTAAAGGAAGCAAAATATAAAGATCCAGTGGTGGCTGCTCTCCAAAGTATTTCAA 240
QY 517 GTGTCTTATCAAGAGGAGATAGCAATTACAGTGGTGAATGGGAAGTCTTGGCCGGCA 576
DB 241 GTGTCTTATCAAGAGGAGATAGCAATTACAGTGGTGAATGGGAAGTCTTGGCCGGCA 300
QY 577 TGGATCTAATATCCAGTTGATGAAGTCAGAGGCTGTATATTTGAAGGCCACCAAGA 636
DB 301 TGGATCTAATATCCAGTTGATGAAGTCAGAGGCTGTATATTTGAAGGCCACCAAGA 360
QY 637 CTCCTCTTTAGAGCATCACCTGTACGTAGTACGTTACGTAATCTCGAGAGGTGACAAG 696
DB 361 CTCCTCTTTAGAGCATCACCTGTACGTAGTACGTTACGTAATCTCGAGAGGTGACAAG 420
QY 697 GCTGACTGACCGTGGCTACTCATTCTTGTGCAATCAGTCAGCAGTGTGACTTCTTTAT 756
DB 421 GCTGACTGACCGTGGCTACTCATTCTTGTGCAATCAGTCAGCAGTGTGACTTCTTTAT 480
QY 757 AAGTAAGTATAGTAACAGAGAAATCCACACTGTGTGTCCTTTTACAAGCTATCAAGTCC 816

```
Db 481 AAGTAAGTAGTAGTAACACAGAGAAATCCACACTGTGTGTCCTTTACAGCTATCAAGTCC 540
Qy 817 TGAAGATGACCACTTGCACAAAGGAAATTTTGGGCCACCAATTTTGGATTTCAGCAGG 876
Db 541 TGAAGATGACCACTTGCACAAAGGAAATTTTGGGCCACCAATTTTGGATTTCAGCAGG 600
Qy 877 TCCTCTTCCCTGACTACTCTCCAGAAATTTTCTCTTTGAAAGTACTACTGGAATTAC 936
Db 601 TCCTCTTCCCTGACTACTCTCCAGAAATTTTCTCTTTGAAAGTACTACTGGAATTAC 660

RESULT 14
AF175225 1292 bp mRNA linear HTC 01-AUG-2003
LOCUS Homo sapiens tissue-type aorta MSTP135 mRNA, complete cds.
AF175225
ACCESSION AF175225.1 GI:33338055
VERSION
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 1292)
Zhao,B., Xu,H.S., Tong,Y.K., Sheng,H., Qin,B.M., Liu,Y.Q., Liu,B.,
Wang,X.Y., Zhang,Q., Song,L., Gao,Y., Zhang,C.L., Ye,J., Ji,X.J.,
Liu,B.H., Lu,H., Chen,J.Z., Cai,M.Q., Zheng,W.Y., Teng,C.Y.,
Liu,Q., Yu,L.T., Lin,J., Gong,Q., Zhang,A.M., Gao,R.L. and Hui,R.T.
Direct Submission
Submitted (04-AUG-1999) Molecular Medicine Center for
Cardiovascular Disease, Cardiovascular Institute, CAMS & PUMC, 167,
Bei Li Shi Lu, Beijing 100037, P.R. China
FEATURES
Location/Qualifiers
1..1292
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="aorta"
142..1623
/codon_start=1
/product="MSTP135"
/protein_id="AA013650.1"
/db_xref="GI:33338056"
/sequence="ATGAAKQNFPGPFWIOVULFTLLQKFLSKLVLLDLHMGCS
TSPHSTGLERILLXLFYIGPKVQLVNNFKGKFFRLNLTLPFLGYVVGIDNRGS
CPRLKGFAPKFGQIEIDQVEGLQYLASRYDFIDLRVGIHGSYGGYLSMAL
MORSIDFRVAGALWLIWIFDYGTIERYMGHPDQNEQYILGSVAMQAEKFPSEPN
RLLHLGLDENVHPAHTSILLSLFLVRAGKPYDLQIYQERHSIRVPESGEHYELHLL
HILQENLGSRIAALKVI"
ORIGIN
Query Match 38.7%; Score 646.4; DB 4; Length 1292;
Best Local Similarity 99.8%; Pred. No. 6.1e-156;
Matches 647; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1021 GGTGTCTATTGTGGGGCCCCAGTCACCTCTCTGGATCTTCTATGATACAGGATACACGGA 1080
Db 645 GGTGTCTATTGTGGGGCCCCAGTCACCTCTCTGGATCTTCTATGATACAGGATACACGGA 704
Qy 1081 AGCTTATATGGGTGTCACCTGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCCAT 1140
Db 705 AGCTTATATGGGTGTCACCTGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCCAT 764
Qy 1141 GCAAGCAGAAAAAGTTCCCTCTGTAACCAAAATCGTTTACTGCTCTTACATGGTTTCTCTGGA 1200
Db 765 GCAAGCAGAAAAAGTTCCCTCTGTAACCAAAATCGTTTACTGCTCTTACATGGTTTCTCTGGA 824
Qy 1201 TGAGAAATGCCATTTTGCACATACAGTATATTACTGAGTTTTTTTGTAGTGGCGCTGGAA 1260
Db 825 TGAGAAATGCCATTTTGCACATACAGTATATTACTGAGTTTTTTTGTAGTGGCGCTGGAA 884
Qy 1261 GCCATATGATTTTACAGATCTATCTCTCAGGAGACACAGCATGAAGATTCCTGGAATCGG 1320
```

```
Db 885 GCCATATGATTTTCAGATCTATCTCTCAGGAGAGACACAGCATAAAGATTCCTGAATCGGG 944
Qy 1321 AGAACATATTGAATCGCATCTTTTGGCACTACCTTCAAGAAAAACCTTGGATCACGTAATTCG 1380
Db 945 AGAACATATTGAATCGCATCTTTTGGCACTACCTTCAAGAAAAACCTTGGATCACGTAATTCG 1004
Qy 1381 TGCTCTAAAAAGTGATATAAATTTTGGACCTGTGTAGAACTCTCTGGTATACACTGGCTATTT 1440
Db 1005 TGCTCTAAAAAGTGATATAAATTTTGGACCTGTGTAGAACTCTCTGGTATACACTGGCTATTT 1064
Qy 1441 AACCAAAATGAGGAGGTTTAAATCAACAGAAAAACACAGAAATTTGATCATCATCTTTCATACC 1500
Db 1065 AACCAAAATGAGGAGGTTTAAATCAACAGAAAAACACAGAAATTTGATCATCATCTTTCATACC 1124
Qy 1501 TGCCATGTAACTACTCTCTGAAAAATAAATGTGGTGCATCGAGGGGTCTACGGTTTGT 1560
Db 1125 TGCCATGTAACTACTCTCTGAAAAATAAATGTGGTGCATCGAGGGGTCTACGGTTTGT 1184
Qy 1561 GGTAGTAATCTAATACCTTAAACCCACATGCTCAAAATCAAAATGATACATATTCCTGAGA 1620
Db 1185 GGTAGTAATCTAATACCTTAAACCCACATGCTCAAAATCAAAATGATACATATTCCTGAGA 1244
Qy 1621 GACCCAGCAATACCATTAAGAAATTTACTTAAAAAARAAAAAARAAAAA 1668
Db 1245 GACCCAGCAATACCATTAAGAAATTTCTTAAAAAARAAAAAARAAAAA 1292

RESULT 15
BU631054 722 bp mRNA linear EST 23-SEP-2002
LOCUS UI-H-FEI-bdt-j-05-0-UI.s1 NCI-CGAP_F51 Homo sapiens cDNA clone
DEFINITION UI-H-FEI-bdt-j-05-0-UI 3', mRNA sequence.
ACCESSION BU631054
VERSION
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 722)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgaaps@mail.nih.gov
Tissue Procurement: James Martin
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, bento-soares@uiowa.edu
Seq primer: M13 FORWARD
POLYA=Yes.
FEATURES
Location/Qualifiers
1..722
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-H-FEI-bdt-j-05-0-UI"
/tissue_type="Cell lines"
/development="Adult"
/lab_host="DH10B (Life Technologies)"
/notes="Organ: Chondrosarcoma; Vector: p7T3-Pac
(pharmacia) with a modified polylinker; Site_1: EcoR I;
Site_2: Not I; NCI CGAP_F51 is a normalized cDNA library
derived from a pool of mRNA obtained from 3 cell lines
from grade II chondrosarcoma tissues. The library was
constructed according to Bonaudo, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
```


synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is CGCTACGGAC. The cell lines were provided by Dr James Martin from the University of Iowa.
TAG TISSUE=Human grade 2 chondrosarcoma cell line pool
TAG_LIB=UI-H-FBI
TAG_SEQ=CGCTACGGAC"

ORIGIN

Query Match		38.6%	Score 644;	DB 5;	Length 722;
Best Local Similarity		100.0%	Pred. No. 2.2e-155;	Indels 0;	Gaps 0;
Matches 644;		Conservative 0;	Mismatches 0;		
Qy	1021	GGTTGCTATTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGA	1080		
Db	644	GGTTGCTATTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGA	585		
Qy	1081	ACGTTATATGGGTACCGTACACGAGATGAACAGGGCTATTACTTAGGATCTGTGGCAT	1140		
Db	584	ACGTTATATGGGTACCGTACACGAGATGAACAGGGCTATTACTTAGGATCTGTGGCAT	525		
Qy	1141	GCAAGCAGAAAAGTTCCTCTGACCAAAATCGTTTACTGTCTTTACATGGTTTCCTGGA	1200		
Db	524	GCAAGCAGAAAAGTTCCTCTGACCAAAATCGTTTACTGTCTTTACATGGTTTCCTGGA	465		
Qy	1201	TGAGAATGTCATTTTGCACATACACAGTATATTACTGAGTTTTTTAGTGGGGCTGAAA	1260		
Db	464	TGAGAATGTCATTTTGCACATACACAGTATATTACTGAGTTTTTTAGTGGGGCTGAAA	405		
Qy	1261	GCCATATGATTTACAGATCTATCCTCAGGAGAGACACAGATTAAGATTCTGAATCGGG	1320		
Db	404	GCCATATGATTTACAGATCTATCCTCAGGAGAGACACAGATTAAGATTCTGAATCGGG	345		
Qy	1321	AGAACATTGAACTGATCTTTTGGCACTACCTTCAAGAAAACCTTGGATCAGCTATTGC	1380		
Db	344	AGAACATTGAACTGATCTTTTGGCACTACCTTCAAGAAAACCTTGGATCAGCTATTGC	285		
Qy	1381	TGCTCTAAAAGTGATATAATTTTGACCTGTGTAGAACTCTCTGGTATACACTGGCTATTT	1440		
Db	284	TGCTCTAAAAGTGATATAATTTTGACCTGTGTAGAACTCTCTGGTATACACTGGCTATTT	225		
Qy	1441	AACCAATGAGGAGGTTTAAATCAACAGAAAACACAGATTTGATCATCATTTTGATACC	1500		
Db	224	AACCAATGAGGAGGTTTAAATCAACAGAAAACACAGATTTGATCATCATTTTGATACC	165		
Qy	1501	TGCCATCTAACATCTACTCTCGAAATAATGTGGTGGCCATGCGGGGTCTACGGTTGT	1560		
Db	164	TGCCATCTAACATCTACTCTCGAAATAATGTGGTGGCCATGCGGGGTCTACGGTTGT	105		
Qy	1561	GGTAGTAATCTTAATACCTTAACCCCATGCTCAAAATCAAAATGATACATATTCCTGAGA	1620		
Db	104	GGTAGTAATCTTAATACCTTAACCCCATGCTCAAAATCAAAATGATACATATTCCTGAGA	45		
Qy	1621	GACCCAGCATATACATTAAGATTACTAAAAAATAAAAAA	1664		
Db	44	GACCCAGCATATACATTAAGATTACTAAAAAATAAAAAA	1		

Search completed: May 4, 2006, 07:56:08
Job time : 6266.09 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 3, 2006, 01:01:26 ; Search time 528.298 Seconds
(without alignments)
5615.673 Million cell updates/sec

Title: US-10-825-632-6
Perfect score: 1669
Sequence: 1 aacaggtagcagaatccta.....aaaaaaaaaaaaaaaaaaaa 1669

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/prodata/1/ina/1 COMB.seq: *
2: /cgn2_6/prodata/1/ina/5 COMB.seq: *
3: /cgn2_6/prodata/1/ina/6A COMB.seq: *
4: /cgn2_6/prodata/1/ina/6B COMB.seq: *
5: /cgn2_6/prodata/1/ina/H COMB.seq: *
6: /cgn2_6/prodata/1/ina/PCTUS COMB.seq: *
7: /cgn2_6/prodata/1/ina/PP COMB.seq: *
8: /cgn2_6/prodata/1/ina/RE COMB.seq: *
9: /cgn2_6/prodata/1/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1669	100.0	1669	3	US-10-070-464-6
2	1648.4	98.8	4523	3	US-09-976-674-8
3	1501	89.9	4876	3	US-09-976-674-20
4	1349.6	80.9	4829	3	US-09-976-674-12
5	1347	80.7	3120	3	US-10-070-464-2
6	1343.6	80.5	4685	3	US-09-976-674-22
7	1337	80.1	2797	3	US-09-976-594-1103
8	1085.8	65.1	2671	3	US-09-976-674-2
9	881	52.8	1197	3	US-10-070-464-4
10	765.2	45.8	4309	3	US-09-976-674-14
11	493.4	29.6	823	3	US-09-280-116-171
12	434.2	26.0	1083	3	US-10-070-464-8
13	313	18.8	2617	3	US-09-976-674-4
14	313	18.8	4076	3	US-09-976-674-32
15	313	18.8	4159	3	US-09-976-674-30
16	313	18.8	4219	3	US-09-976-674-28
17	313	18.8	4302	3	US-09-976-674-24
18	297	17.8	308	3	US-09-621-976-1787
19	252.6	15.1	612	3	US-09-392-184-31
20	250.6	15.0	4037	3	US-09-976-674-40
21	250.6	15.0	4120	3	US-09-976-674-38
22	250.6	15.0	4180	3	US-09-976-674-36
23	250.6	15.0	4263	3	US-09-976-674-34
24	123.2	7.4	2411	3	US-09-976-674-26

25	89	5.3	89	3	US-09-513-999C-16433	Sequence 16433, A
26	74.8	4.5	936	3	US-09-270-767-12940	Sequence 12940, A
27	45.6	2.7	2924	3	US-09-023-655-1026	Sequence 1026, Ap
28	45.6	2.7	2924	6	PCT-US93-07923-1	Sequence 1, Appli
29	45.6	2.7	3407	3	US-10-002-593-5	Sequence 5, Appli
30	45.6	2.7	3407	3	US-09-949-016-275	Sequence 275, App
31	45.6	2.7	3407	3	US-10-423-714-5	Sequence 5, Appli
32	45.6	2.7	3419	3	US-09-949-016-4579	Sequence 4579, Ap
33	45.4	2.7	1023	3	US-09-270-767-13726	Sequence 13726, A
34	42.8	2.6	480	3	US-09-270-767-684	Sequence 684, App
35	42.8	2.6	480	3	US-09-270-767-15966	Sequence 15966, A
36	42.4	2.5	972	2	US-07-915-934-1	Sequence 1, Appli
37	42.4	2.5	972	2	US-08-325-743-1	Sequence 22, Appli
38	41.2	2.5	1141	3	US-09-806-708B-22	Sequence 11817, A
39	40.8	2.4	124700	3	US-09-949-016-11817	Sequence 13439, A
40	40.8	2.4	124701	3	US-09-949-016-15439	Sequence 22, Appli
41	40.6	2.4	1141	3	US-09-806-708B-22	Sequence 2813, Ap
42	40.2	2.4	832	3	US-09-621-976-2813	Sequence 13081, A
43	40	2.4	708	3	US-09-270-767-13081	Sequence 9, Appli
44	39.8	2.4	1405	3	US-09-244-111-9	Sequence 64, Appli
45	39.2	2.3	1474	3	US-08-821-994-64	

ALIGNMENTS

RESULT 1
US-10-070-464-6
; Sequence 6, Application US/10070464
; Patent No. 6881564
; GENERAL INFORMATION:
; APPLICANT: ABBOTT, Catherine Anne
; APPLICANT: GORRELL, Mark Douglas
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES
; FILE REFERENCE: GH-007
; CURRENT APPLICATION NUMBER: US/10/070,464
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: PCT/AU00/01085
; PRIOR FILING DATE: 2000-09-11
; PRIOR APPLICATION NUMBER: AU PQ5709
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: AU PQ2762
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1669
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-10-070-464-6

Query Match 100.0%; Score 1669; DB 3; Length 1669;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1669; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	AACAGGTACAGCAAACTCTTAAGTCTCTTTAAGTGTCTAGAGTGTCTAGTGTCTGCTGA	60
Db	1	AACAGGTACAGCAAACTCTTAAGTCTCTTTAAGTGTCTAGAGTGTCTAGTGTCTGCTGA	60
Qy	61	AGGAAGGATCATAGATGTCTAGATTAAGCACTAATTCACCTTTGAGATTCCTATTGA	120
Db	61	AGGAAGGATCATAGATGTCTAGATTAAGCACTAATTCACCTTTGAGATTCCTATTGA	120
Qy	121	AGGAGTTGATATATTGCCAGAGCTGCTGAGTCTCTCTGAGGAAATATGCTTGTCTCAT	180
Db	121	AGGAGTTGATATATTGCCAGAGCTGCTGAGTCTCTCTGAGGAAATATGCTTGTCTCAT	180
Qy	181	CCTACTAGATCGCTCCAGACTCGCCTACAGATAGTGTGTGATCTCCACCTGAATTTAT	240
Db	181	CCTACTAGATCGCTCCAGACTCGCCTACAGATAGTGTGTGATCTCCACCTGAATTTAT	240
Qy	241	CCCAGTGAAGATGATGTTTATGAAAGGAGAGACCTCATGAGTCAGTGCCTGATCTGT	300
Db	241	CCCAGTGAAGATGATGTTTATGAAAGGAGAGACCTCATGAGTCAGTGCCTGATCTGT	300

```
Db 241 CCCAGTAGAAGATGATGTTTATGGAAGGCAGAGACTCATTGAGTCAGTGCCTGATTCTGT 300
Qy 301 GAGCCCACTAATTTATCTATGAGAAACAACACACATCTGGATAAATATCCATGACATCTT 360
Db 301 GAGCCCACTAATTTATCTATGAGAAACAACACACATCTGGATAAATATCCATGACATCTT 360
Qy 361 TCATGTTTTTCCCAAGTCACGAGAGGAAATTCAGTTTTATTTTTCGCTCTGATGCAA 420
Db 361 TCATGTTTTTCCCAAGTCACGAGAGGAAATTCAGTTTTATTTTTCGCTCTGATGCAA 420
Qy 421 AACAGGTTCCGTCATTTATACAAATTTACATCTATTTTAAAGAAAGCAAAATATAACG 480
Db 421 AACAGGTTCCGTCATTTATACAAATTTACATCTATTTTAAAGAAAGCAAAATATAACG 480
Qy 481 ATCCAGTGTGGCTGCCCTGCCCAAGTGATTTCAAGTGTCTCTATCAAAAGAGGAGATAGC 540
Db 481 ATCCAGTGTGGCTGCCCTGCCCAAGTGATTTCAAGTGTCTCTATCAAAAGAGGAGATAGC 540
Qy 541 AATTACAGTGTGGAATGGAAAGTTCTTGGCGGCATGGATCTAATATCCAAGTTGATGA 600
Db 541 AATTACAGTGTGGAATGGAAAGTTCTTGGCGGCATGGATCTAATATCCAAGTTGATGA 600
Qy 601 AGTCAGAAGCTGGTATATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACCTGTA 660
Db 601 AGTCAGAAGCTGGTATATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACCTGTA 660
Qy 661 CGTAGTCAGTTAGTAAATCTCGAGAGAGTGACAAGGCTGACTGACCGTGGCTACTCACA 720
Db 661 CGTAGTCAGTTAGTAAATCTCGAGAGAGTGACAAGGCTGACTGACCGTGGCTACTCACA 720
Qy 721 TTCTCTGCTGCATCAGTCAGCAGCTGACTTCTTTTAAGTAAGTATAGTAAACCAAGAA 780
Db 721 TTCTCTGCTGCATCAGTCAGCAGCTGACTTCTTTTAAGTAAGTATAGTAAACCAAGAA 780
Qy 781 TCACACTGTGTGCTTCAAGCTATCAAGTCTGGAAGTGAACCACTTGGAACCA 840
Db 781 TCACACTGTGTGCTTCAAGCTATCAAGTCTGGAAGTGAACCACTTGGAACCA 840
Qy 841 AAGGAATTTTGGGACACATTTTGGATTACAGAGGCTCTCTCTCTGACTATATCTCTCC 900
Db 841 AAGGAATTTTGGGACACATTTTGGATTACAGAGGCTCTCTCTCTGACTATATCTCTCC 900
Qy 901 AGAAATTTCTCTTTTGAAGTACTACTGGAATTTACATTTGATGGATGCTCTACAGCC 960
Db 901 AGAAATTTCTCTTTTGAAGTACTACTGGAATTTACATTTGATGGATGCTCTACAGCC 960
Qy 961 TCATGATCTACAGCTGGAAAGAAATATCTTACTGTCTGTTTCAATATATGGTGGTCTCA 1020
Db 961 TCATGATCTACAGCTGGAAAGAAATATCTTACTGTCTGTTTCAATATATGGTGGTCTCA 1020
Qy 1021 GGTGCTATTTGCTGGGGCCCGAGTCACCTCTGTGGATCTTCTATGATACAGGATACACGGA 1080
Db 1021 GGTGCTATTTGCTGGGGCCCGAGTCACCTCTGTGGATCTTCTATGATACAGGATACACGGA 1080
Qy 1081 ACGTTATATGGGTCAACCTGACCAAGATGAACAGGGCTATTCTTAGGATCTGTGGCCAT 1140
Db 1081 ACGTTATATGGGTCAACCTGACCAAGATGAACAGGGCTATTCTTAGGATCTGTGGCCAT 1140
Qy 1141 GCAAGCAGAAAGTTCCCTCTGACCAAAATCGTTTACTGCTCTTACATGGTTTCTGGA 1200
Db 1141 GCAAGCAGAAAGTTCCCTCTGACCAAAATCGTTTACTGCTCTTACATGGTTTCTGGA 1200
Qy 1201 TCAGATGTCCTTTTGCATACAGTATATTTACTGAGTTTTTTTATGAGGGCTGGAAA 1260
Db 1201 TCAGATGTCCTTTTGCATACAGTATATTTACTGAGTTTTTTTATGAGGGCTGGAAA 1260
Qy 1261 GCCATATGATTTACAGATCTATCTCAGGAGACACAGCATTAAGAGTTCTCGAATCGGG 1320
Db 1261 GCCATATGATTTACAGATCTATCTCAGGAGACACAGCATTAAGAGTTCTCGAATCGGG 1320
Qy 1321 AGAACTATTAAGCTGCATCTTTTGGCACTACCTTCAAGAAACCTTTGGATCACGATTGCG 1380
Db 1321 AGAACTATTAAGCTGCATCTTTTGGCACTACCTTCAAGAAACCTTTGGATCACGATTGCG 1380
```

```
Qy 1381 TGCTCTAAAAGTGATATATATTTTGGACCTGTGTAGAACTCTCTGGTATATACATCGGCTATTT 1440
Db 1381 TGCTCTAAAAGTGATATATATTTTGGACCTGTGTAGAACTCTCTGGTATATACATCGGCTATTT 1440
Qy 1441 AACCAATATGAGGAGGTTTTTAAATCAACAGAAAAACAGAAATTTGATCATCATCTTTGATACC 1500
Db 1441 AACCAATATGAGGAGGTTTTTAAATCAACAGAAAAACAGAAATTTGATCATCATCTTTGATACC 1500
Qy 1501 TGCCATGTAACTACTCTCTGAAAAATAAATCTGGTGCATGCGAGGGTCTACGGTTTGT 1560
Db 1501 TGCCATGTAACTACTCTCTGAAAAATAAATCTGGTGCATGCGAGGGTCTACGGTTTGT 1560
Qy 1561 GGTAGTAAATCTAATACCTTTAAACCCACATGCTCAAAATCAAAATGATACATATTTCTGAGA 1620
Db 1561 GGTAGTAAATCTAATACCTTTAAACCCACATGCTCAAAATCAAAATGATACATATTTCTGAGA 1620
Qy 1621 GACCCAGCAATACCAATAAGAAATTTACTTAAAAAATAAATTTTAAAAAATAAATTTTAAAA 1669
Db 1621 GACCCAGCAATACCAATAAGAAATTTACTTAAAAAATAAATTTTAAAAAATAAATTTTAAAA 1669
```

RESULT 2

```
US-09-976-674-8
; Sequence 8, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 4523
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-976-674-8
```

```
Query Match 98.8%; Score 1648.4; DB 3; Length 4523;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1663; Conservative 0; Mismatches 1; Indels 4; Gaps 1;
```

```
Qy 1 AACAGGTACAGCAAAATCCTAAAGTCACCTTTAAAGATGTCAGAAATATGATGATGCTGA 60
Db 1164 AACAGGTACAGCAAAATCCTAAAGTCACCTTTAAAGATGTCAGAAATATGATGATGCTGA 1223
Qy 61 AGGAAGGATCATAGATGTCATAGATAAGGAACCTAAATTCACCTTTTGAGATTTCTATTGA 120
Db 1224 AGGAAGGATCATAGATGTCATAGATAAGGAACCTAAATTCACCTTTTGAGATTTCTATTGA 1283
Qy 121 AGGAGTTGANTATATTTGCCAGAGCTGGATGACTCTCGAGGAAAAATATGCTTGGTCCAT 180
Db 1284 AGGAGTTGANTATATTTGCCAGAGCTGGATGACTCTCGAGGAAAAATATGCTTGGTCCAT 1343
Qy 181 CCTACTAGATCGCTCCCAGACTCGCCTACAGATAGTGTGATCTCACCTCAATTTATTAT 240
Db 1344 CCTACTAGATCGCTCCCAGACTCGCCTACAGATAGTGTGATCTCACCTCAATTTATTAT 1403
Qy 241 CCCAGTAGAAGATGATGTTTATGGAAAGCGAGAGACTCATTTGAGTCAGTCGCTGATTTCTGT 300
Db 1404 CCCAGTAGAAGATGATGTTTATGGAAAGCGAGAGACTCATTTGAGTCAGTCGCTGATTTCTGT 1463
Qy 301 GACGCCACTAATTTATCTATGAAGAAAAACAACAGACATCTGGATAAATATCCATGACATCTT 360
Db 1464 GACGCCACTAATTTATCTATGAAGAAAAACAACAGACATCTGGATAAATATCCATGACATCTT 1523
```

QY 361 TCATGTTTTCCTCCCAAGTACGAGAGGAAATTTGAGTTTATTTTTTGCTCTGATGCAA 420
Db |||||
1524 TCATGTTTTCCTCCCAAGTACGAGAGGAAATTTGAGTTTATTTTTTGCTCTGATGCAA 1583
QY 421 AACAGGTTTCCTGCTATTTATACAAAATTTACATCTATTTTAAAGGAAAGCAATATAAACG 480
Db |||||
1584 AACAGGTTTCCTGCTATTTATACAAAATTTACATCTATTTTAAAGGAAAGCAATATAAACG 1643
QY 481 ATCCAGTGGTGGGCTGCTGCTCCCAAGTATTTCAAGTGTCTCTATCAAGAGGAGATAGC 540
Db |||||
1644 ATCCAGTGGTGGGCTGCTGCTCCCAAGTATTTCAAGTGTCTCTATCAAGAGGAGATAGC 1703
QY 541 AATTACCAAGTGGTGAATGGGAGGTTCTTGGCCGGCATGGATCTAATATCCAAGTTGATGA 600
Db |||||
1704 AATTACCAAGTGGTGAATGGGAGGTTCTTGGCCGGCATGGATCTAATATCCAAGTTGATGA 1763
QY 601 AGTCAGAAGGCTGGTATATTTTGAAGGCACCAAGAGCTCCCTTTAGAGCATCACCTGTA 660
Db |||||
1764 AGTCAGAAGGCTGGTATATTTTGAAGGCACCAAGAGCTCCCTTTAGAGCATCACCTGTA 1823
QY 661 CGTAGTCAGTTACGTAATCTCGAGAGGTGACAAGGCTGACTGACCGTGGCTACTACA 720
Db |||||
1824 CGTAGTCAGTTACGTAATCTCGAGAGGTGACAAGGCTGACTGACCGTGGCTACTACA 1883
QY 721 TTCTTTGCTGCATCAGTCAGCACTGTGACTTCTTTATAGTAAGTATAGTAACACGAGAA 780
Db |||||
1884 TTCTTTGCTGCATCAGTCAGCACTGTGACTTCTTTATAGTAAGTATAGTAACACGAGAA 1943
QY 781 TCACACATGTGTGCTCCCTTTTACAAGTATCAAGTCTCGAAGATGACCCCACTGCAAAAC 840
Db |||||
1944 TCACACATGTGTGCTCCCTTTTACAAGTATCAAGTCTCGAAGATGACCCCACTGCAAAAC 2003
QY 841 AAAGGAATTTTGGGCCACCATTTTGGATTACAGAGGTCCTCTCTGACTATACCTCC 900
Db |||||
2004 AAAGGAATTTTGGGCCACCATTTTGGATTACAGAGGTCCTCTCTGACTATACCTCC 2063
QY 901 AGAATTTTCTCTTTTGAAGTACTTCTGGAATTTACATTTGTATGGATGCTCTTACAAGCC 960
Db |||||
2064 AGAATTTTCTCTTTTGAAGTACTTCTGGAATTTACATTTGTATGGATGCTCTTACAAGCC 2123
QY 961 TCATGATCTACAGCTCGAAGAAATATCTTACTGTCTGTCTCATATATGTTGGTCTCTCA 1020
Db |||||
2124 TCAATATCTACAGCTCGAAGAAATATCTTACTGTCTGTCTCATATATGTTGGTCTCTCA 2179
QY 1021 GGTGTCTATTTGCTGGGGCCCCAGTCACTCTGTGGAATCTTCTATGATACAGGATACAGGA 1080
Db |||||
2180 GGTGTCTATTTGCTGGGGCCCCAGTCACTCTGTGGAATCTTCTATGATACAGGATACAGGA 2239
QY 1081 AGTTATATGGGTACCCCTGACCAAGATGAACAGGGCTATTTACTTAGGATCTGTGGCCAT 1140
Db |||||
2240 AGTTATATGGGTACCCCTGACCAAGATGAACAGGGCTATTTACTTAGGATCTGTGGCCAT 2299
QY 1141 GCAAGCAGAAAGTTCCCTCTGAAACCAATCGTTTACTGTCTTTACATGTTTCTCTGGA 1200
Db |||||
2300 GCAAGCAGAAAGTTCCCTCTGAAACCAATCGTTTACTGTCTTTACATGTTTCTGGA 2359
QY 1201 TGGAATGTCCATTTTGGACATACCAAGTATATTTACTGAGTTTCTTTAGTGGGCTGGAAA 1260
Db |||||
2360 TGGAATGTCCATTTTGGACATACCAAGTATATTTACTGAGTTTCTTTAGTGGGCTGGAAA 2419
QY 1261 GCCATATGATTTACAGATCTATCTCAGGAGAGACACAGATTAAGATTTCTGATCGGG 1320
Db |||||
2420 GCCATATGATTTACAGATCTATCTCAGGAGAGACACAGATTAAGATTTCTGATCGGG 2479
QY 1321 AGAATATGAACTGTGATCTTTTGCACTACCTTCAAGAAACCTTGGATCAGCTATTC 1380
Db |||||
2480 AGAATATGAACTGTGATCTTTTGCACTACCTTCAAGAAACCTTGGATCAGCTATTC 2539
QY 1381 TGCTCTAAAAGTGTATATTTTGGACCTGTGTAGAACTCTCTGGTATACACTGGCTATTT 1440
Db |||||
2540 TGCTCTAAAAGTGTATATTTTGGACCTGTGTAGAACTCTCTGGTATACACTGGCTATTT 2599
QY 1441 AACCBAATGAGGAGTTTAAATCAACAGAAACACAGAAATGATCATCATTTTGTATACC 1500

Db 2600 AACCAATGAGGAGTTTAAATCAACAGAAACACAGAAATGATCATCATTTTGTATACC 2659
QY 1501 TGCCATGTAAACATCTACTCTCTGAAATAAATGTGGTCCCATGACGGGTCTACGGTTTGT 1560
Db 2660 TGCCATGTAAACATCTACTCTCTGAAATAAATGTGGTCCCATGACGGGTCTACGGTTTGT 2719
QY 1561 GGTAGTAAATCTAAATACCTTTAAACCCACATGCTCAAAATCAAAATGATACATATTTCTCTGAGA 1620
Db 2720 GGTAGTAAATCTAAATACCTTTAAACCCACATGCTCAAAATCAAAATGATACATATTTCTCTGAGA 2779
QY 1621 GACCCAGCAATACCATTAAGAAATTAATAAATAAATAAATAAATAAATAAATAAATAA 1668
Db 2780 GACCCAGCAATACCATTAAGAAATTAATAAATAAATAAATAAATAAATAAATAAATAA 2827

RESULT 3
US-09-976-674-20
; Sequence 20, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: Patent In version 3.1
; SEQ ID NO 20
; LENGTH: 4676
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-976-674-20

Query Match 89.9%; Score 1501; DB 3; Length 4676;
Best Local Similarity 91.5%; Pred. No. 0;
Matches 1663; Conservative 0; Mismatches 5; Indels 149; Gaps 1;

QY 1 AACAGGTACAGCAAAATCCTAAAGTCACCTTTTAAAGATGTCAGAAATATATGATGCTGA 60
Db 1164 AACAGGTACAGCAAAATCCTAAAGTCACCTTTTAAAGATGTCAGAAATATATGATGCTGA 1223
QY 61 AGAAGGATCATAGATGTCATAGATAAGTAAGTAATTCACCTTTTGAGATTCATTTGA 120
Db 1224 AGAAGGATCATAGATGTCATAGATAAGTAAGTAATTCACCTTTTGAGATTCATTTGA 1283
QY 121 AGAGTTGATATATATGTCAGAGCTGGATGGACTCTCAGGAGAAATATGCTGGTCCAT 180
Db 1284 AGAGTTGATATATATGTCAGAGCTGGATGGACTCTCAGGAGAAATATGCTGGTCCAT 1343
QY 181 CCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAATTTAT 240
Db 1344 CCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAATTTAT 1403
QY 241 CCCAGTAGAAGATGATGTTTATGAAAGGACAGATCAATGAGTCAGTCGCTGATCTGT 300
Db 1404 CCCAGTAGAAGATGATGTTTATGAAAGGACAGATCAATGAGTCAGTCGCTGATCTGT 1463
QY 301 GAGCCACTAATTTATCTATGAAGAAACACAGACATCTGGATAAATATCCATGACATCTT 360
Db 1464 GAGCCACTAATTTATCTATGAAGAAACACAGACATCTGGATAAATATCCATGACATCTT 1523
QY 361 TCATGTTTTTCCCCAAAGTCAGAGGAAATTCAGTTTATTTTGGCTCTGTAATGCAA 420
Db 1524 TCATGTTTTTCCCCAAAGTCAGAGGAAATTCAGTTTATTTTGGCTCTGTAATGCAA 1583
QY 421 AACAGGTTTCGTCATTTTATACAAATTTACATCTATTTTAAAGGAAACCAATATAAACG 480

Db 1584 AACAGGTTCCGTCATTTATACAAATTTACATCTATTTTAAAGGAAACCAATATATAACG 1643
QY 481 ATCCAGTGGTGGCTGCTCCCTGCTCCAAAGTGAATTTCAAGTGTCTCTATCAAAAGAGGAGATAGC 540
Db 1644 ATCCAGTGGTGGCTGCTCCCTGCTCCAAAGTGAATTTCAAGTGTCTCTATCAAAAGAGGAGATAGC 1703
QY 541 AATTACAGTGGTGAATGGGAAGTTCTTTGGCCGGGATGGATCTAATATCCAAGTTGATGA 600
Db 1704 AATTACAGTGGTGAATGGGAAGTTCTTTGGCCGGGATGGATCTAATATCCAAGTTGATGA 1763
QY 601 AGTCAGAAGGCTGGTATATTTTGAAGGACCAAAAGACTCCCTTTAGAGCATCACCTGTGA 660
Db 1764 AGTCAGAAGGCTGGTATATTTTGAAGGACCAAAAGACTCCCTTTAGAGCATCACCTGTGA 1823
QY 661 CGTAGTCAGTTACGTAAATCCTGGAGAGGTGACAGGCTGACCGTGGCTACTCACA 720
Db 1824 CGTAGTCAGTTACGTAAATCCTGGAGAGGTGACAGGCTGACCGTGGCTACTCACA 1883
QY 721 TTCTTCTGTCATCAGTCAGCACTGTGACTTCTTTAAGTAAGTATAGTAAACAGAGAA 780
Db 1884 TTCTTCTGTCATCAGTCAGCACTGTGACTTCTTTAAGTAAGTATAGTAAACAGAGAA 1943
QY 781 TCCACACTGTGTCCCTTTACAGCTATCAAGTCTCTGAAGTGAACCAACTTGCAAAAC 840
Db 1944 TCCACACTGTGTCCCTTTACAGCTATCAAGTCTCTGAAGTGAACCAACTTGCAAAAC 2003
QY 841 AAGGAAATTTGGGACACATTTGGATTTCAGAGGTCCTTCTCTGACTATACCTCTCC 900
Db 2004 AAGGAAATTTGGGACACATTTGGATTTCAGAGGTCCTTCTCTGACTATACCTCTCC 2063
QY 901 AGNAATTTCTCTTTGAAAGTACTACTGGAATTTACATTTGGAATGCTCTACAGACC 960
Db 2064 AGNAATTTCTCTTTGAAAGTACTACTGGAATTTACATTTGGAATGCTCTACAGACC 2123
QY 961 TCATGATTCACAGCTGGAAGAAATATCTCTGCTGTCTCTCATATATGGTGGTCT --- 1016
Db 2124 TCATGATTCACAGCTGGAAGAAATATCTCTGCTGTCTCTCATATATGGTGGTCTGGGT 2183
QY 1017 ----- 1016
Db 2184 CAATAGAAATTCAGCATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTC 2243
QY 1017 ----- 1016
Db 2244 ATTGACTTAGATCGTGTGGGCATCCAGGCTGGTCTCTATGGAGATACCTCTCCCTGATG 2303
QY 1017 -----CTCAGTTGCTATTGCTGGGGCCCGCTCACTCTG 1051
Db 2304 GCATTAATGCAGAGTCAGATATCTTCAGGGTTGCTATTGCTGGGGCCCGCTCACTCTG 2363
QY 1052 TGGATCTTCTATGATACAGGATACACGGAACGTTATATGSGTCACCCCTGACCGAATGAA 1111
Db 2364 TGGATCTTCTATGATACAGGATACACGGAACGTTATATGSGTCACCCCTGACCGAATGAA 2423
QY 1112 CAGGCTATTACTTAGGATCTGTGGCCATGCAAGCAAAAGTTCCCTCTGGAACCAAT 1171
Db 2424 CAGGCTATTACTTAGGATCTGTGGCCATGCAAGCAAAAGTTCCCTCTGGAACCAAT 2483
QY 1172 CGTTTACTGCTTTACATGTTTCTGGATGAGAAATGTCATTTTGGCATACACCAATATA 1231
Db 2484 CGTTTACTGCTTTACATGTTTCTGGATGAGAAATGTCATTTTGGCATACACCAATATA 2543
QY 1232 TTACTGATTTTACTGAGGCTGGAAGCCATATGATTTACAGATCTATCTCTCAGGAG 1291
Db 2544 TTACTGATTTTACTGAGGCTGGAAGCCATATGATTTACAGATCTATCTCTCAGGAG 2603
QY 1292 AGACACAGCATAGAGTTCTGATCGGGAGAACATTTATGAACTGCATCTTTTGCCTAC 1351
Db 2604 AGACACAGCATAGAGTTCTGATCGGGAGAACATTTATGAACTGCATCTTTTGCCTAC 2663
QY 1352 CTTCAAGAAACCTTGGATCAGCTATTTGCTGTCTTAAAGTGATATAATTTTGACCTGTG 1411
Db 2664 CTTCAAGAAACCTTGGATCAGCTATTTGCTGTCTTAAAGTGATATAATTTTGACCTGTG 2723

RESULT 4

US-09-976-674-12 ; Sequence 12, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; PRIORITY FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 4829
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-976-674-12

Query Match 80.9%; Score 1349.6; DB 3; Length 4829;
Best Local Similarity 84.5%; Pred. No. 0;
Matches 1664; Conservative 0; Mismatches 4; Indels 302; Gaps 1;

QY 1 AACAGGTACAGAAATCTCTAAAGTCACCTTTAAAGATGTCAGAAATATGATGATGCTGA 60
Db 1164 AACAGGTACAGAAATCTCTAAAGTCACCTTTAAAGATGTCAGAAATATGATGATGCTGA 1223
QY 61 AGGAAGGATCATAGATGTCTAGATAAGGAACCTAATTCACCTTTTGAGATTTCTATTGA 120
Db 1224 AGGAAGGATCATAGATGTCTAGATAAGGAACCTAATTCACCTTTTGAGATTTCTATTGA 1283
QY 121 AGGATTTGAATATATTGCCAGAGCTGGATGGAATCTCTGAGGGAATAATATGCTTGGTCCAT 180
Db 1284 AGGATTTGAATATATTGCCAGAGCTGGATGGAATCTCTGAGGGAATAATATGCTTGGTCCAT 1343
QY 181 CCTACTAGATCGCTCCAGAGCTCGCTTACAGATAGTGTGATCTCACCTCCCTGATTTAT 240
Db 1344 CCTACTAGATCGCTCCAGAGCTCGCTTACAGATAGTGTGATCTCACCTCCCTGATTTAT 1403
QY 241 CCCAGTGAAGATGATGTTATGAAAGGACAGAGACTTTGAGTCAGTCCCTGATTCGT 300
Db 1404 CCCAGTGAAGATGATGTTATGAAAGGACAGAGACTTTGAGTCAGTCCCTGATTCGT 1463
QY 301 GACGCCACTAATTTATCTATGAAGAAACACAGACATCTGGATAAATATCCATGACATCTT 360
Db 1464 GACGCCACTAATTTATCTATGAAGAAACACAGACATCTGGATAAATATCCATGACATCTT 1523

Qy	361	TCATGTTTTTCCCAAGTCA	CGAAGAGGAATTGAGTTTAT	TTTTTGGCTCTGAATGCAA	420
Db	1524	TCATGTTTTTCCCAAGTCA	CGAAGAGGAATTGAGTTTAT	TTTTTGGCTCTGAATGCAA	1583
Qy	421	AACAGGTTTCGGTCAATTA	CAAAATACATCTATTTTAA	AGGAAGCAAAATATAAAG	480
Db	1584	AACAGGTTTCGGTCAATTA	CAAAATACATCTATTTTAA	AGGAAGCAAAATATAAAG	1643
Qy	481	ATCCAGTGGTGGCTGCCTG	CTCCAAAGTGATTTCAAGT	GTCTATCAAAAGGAGATAG	540
Db	1644	ATCCAGTGGTGGCTGCCTG	CTCCAAAGTGATTTCAAGT	GTCTATCAAAAGGAGATAG	1703
Qy	541	AAATTACAGTGGTGAATGG	AAAGTTCTTGGCCGGCATGG	ATCTAATATCCAAGTTGATG	600
Db	1704	AAATTACAGTGGTGAATGG	AAAGTTCTTGGCCGGCATGG	ATCTAATATCCAAGTTGATG	1763
Qy	601	AGTCAGAGGCTGGTATAT	TTTTGAAGGCAACCAAGAC	TCCCTTTAGAGCATCACCTGT	660
Db	1764	AGTCAGAGGCTGGTATAT	TTTTGAAGGCAACCAAGAC	TCCCTTTAGAGCATCACCTGT	1823
Qy	661	CGTAGTCAGTTACGTAAT	CCTGGAGAGTGACAAGGCT	GTGACCGTGGCTACTACA	720
Db	1824	CGTAGTCAGTTACGTAAT	CCTGGAGAGTGACAAGGCT	GTGACCGTGGCTACTACA	1883
Qy	721	TTCTTGCTGCATCAGTCAG	CACCTGTGACTTCTTTAAG	TATAGTATAGTACCAAGAG	780
Db	1884	TTCTTGCTGCATCAGTCAG	CACCTGTGACTTCTTTAAG	TATAGTATAGTACCAAGAG	1943
Qy	781	TCCACACTGTGTGCTCTTT	TACAAGCTATCAAGTCTCA	AGATGACCCCACTTGCAAA	840
Db	1944	TCCACACTGTGTGCTCTTT	TACAAGCTATCAAGTCTCA	AGATGACCCCACTTGCAAA	2003
Qy	841	AAAGGAATTTTGGGCGCAC	CAATTTGGGATTCAGAGGT	CTCTCTGCTGACTATATCT	900
Db	2004	AAAGGAATTTTGGGCGCAC	CAATTTGGGATTCAGAGGT	CTCTCTGCTGACTATATCT	2063
Qy	901	AGAAATTTTCTCTTTTGA	AGTACTACTGGATTTACAT	TGTATGGGATGCTCTACA	960
Db	2064	AGAAATTTTCTCTTTTGA	AGTACTACTGGATTTACAT	TGTATGGGATGCTCTACA	2123
Qy	961	TCATGATCTACAGCCTGG	AAAGAAATATCTCTGCTG	CTGTCTATATATGGTGGCT	1020
Db	2124	TCATGATCTACAGCCTGG	AAAGAAATATCTCTGCTG	CTGTCTATATATGGTGGCT	2183
Qy	1021	-----	-----	-----	1020
Db	2184	CAGGTGCAGTTGGTGAAT	AAATCGGTTTAAAGGAGT	CAAGTATTTCCGCTTGAAT	2243
Qy	1021	-----	-----	-----	1020
Db	2244	GCCTCTTAGTGTATGTGG	TGTAGTGATAGACAAGGG	GATCCTGTCCCGAGGGCTT	2303
Qy	1021	-----	-----	-----	1020
Db	2304	AAATTTGAAGGCGCCTTT	AAATATAAAATGGGTCAA	ATAGAAATGACGATCAGGT	2363
Qy	1021	-----	-----	-----	1020
Db	2364	GGACTCCAATATCTAGCT	TCTCGATATGATTTTCTAT	TGACTTAGATCGTGTGGGC	2423
Qy	1021	-----	-----	-----	1020
Db	2424	GGCTGGTCTATGAGGAT	ACCTCTCCCTGATGGCAT	TAAATGACAGGTGAGATAT	2483
Qy	1021	--GGTTGCTATTGCTGGG	CGCCAGTCACTCTGTGGAT	CTCTTCTATGATACAGGAT	1078
Db	2484	AGGGTTGCTATTGCTGGG	CGCCAGTCACTCTGTGGAT	CTCTTCTATGATACAGGAT	2543
Qy	1079	GAACGTTATATGGGTGAC	CCCTGACCAAGTATGAC	AGGGCTATTACTTAGGAT	1138
Db	2544	GAACGTTATATGGGTGAC	CCCTGACCAAGTATGAC	AGGGCTATTACTTAGGAT	2603
Qy	1139	ATCAAGCAGAAAGTTCC	CTCTGAAACCAATCGTTT	ACTGCTCTTACATGTTCTG	1198

Db	2604	ATGCAAGCAGAAAGTTCC	CTCTGAAACCAATCGTTT	TACTGCTCTTACATGGTT	2663
Qy	1199	GATCAGAAATCTCCATTT	TGACATACCACTATATTAC	TACTAGTTTATGAGGGCTG	1258
Db	2664	GATCAGAAATCTCCATTT	TGACATACCACTATATTAC	TACTAGTTTATGAGGGCTG	2723
Qy	1259	AAGCCATATGATTTACAG	ATCTATCTCAGGAGAGAC	CAGCATAGAGTTCTCGAAT	1318
Db	2724	AAGCCATATGATTTACAG	ATCTATCTCAGGAGAGAC	CAGCATAGAGTTCTCGAAT	2783
Qy	1319	GGAAACATTTGAACCTGC	ATCTTTTGCACTACCTTCA	AGAAAACTTTGGATCAGT	1378
Db	2784	GGAAACATTTGAACCTGC	ATCTTTTGCACTACCTTCA	AGAAAACTTTGGATCAGT	2843
Qy	1379	GCTGCTCTAAAAGTGAT	ATAATTTTGACCTGTGTAG	AACTCTCTGGTATACATG	1438
Db	2844	GCTGCTCTAAAAGTGAT	ATAATTTTGACCTGTGTAG	AACTCTCTGGTATACATG	2903
Qy	1439	TTAACCAAAATGAGGAG	TTTTAATCAACAGAAAA	ACAGAAATTTGATCATCA	1498
Db	2904	TTAACCAAAATGAGGAG	TTTTAATCAACAGAAAA	ACAGAAATTTGATCATCA	2963
Qy	1499	CCTGCCATGTAAACATCT	ACTCTCTGAAAAATAAAT	GTGGCCATGAGGGTTT	1558
Db	2964	CCTGCCATGTAAACATCT	ACTCTCTGAAAAATAAAT	GTGGCCATGAGGGTTT	3023
Qy	1559	GTGGTAGTAACTTAATAC	CTTAACCCCATGCTCAAA	ATCAAAATGATACATATT	1618
Db	3024	GTGGTAGTAACTTAATAC	CTTAACCCCATGCTCAAA	ATCAAAATGATACATATT	3083
Qy	1619	GAGACCCAGCAATACCAT	TAAGAAATTTACTAAAA	AAAAAAAAAAAAAAAAAAAA	1668
Db	3084	GAGACCCAGCAATACCAT	TAAGAAATTTACTAAAA	AAAAAAAAAAAAAAAAAAAA	3133

RESULT 5

US-10-070-464-2
; Sequence 2, Application US/10070464
; Patent No. 6881564
; GENERAL INFORMATION:
; APPLICANT: ABBOTT, Catherine Anne
; APPLICANT: GORRELL, Mark Douglas
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES
; FILE REFERENCE: GH-007
; CURRENT APPLICATION NUMBER: US/10/070,464
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: PCT/AU00/01085
; PRIOR FILING DATE: 2000-09-11
; PRIOR APPLICATION NUMBER: AU PQ5709
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: AU PQ2762
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 3120
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-10-070-464-2

Query Match		80.7%;	Score 1347;	DB 3;	Length 3120;
Best Local Similarity		84.7%;	Pred. No. 0;		
Matches 1657;		Conservative 0;	Mismatches 0;	Indels 300;	Gaps 1;
Qy	1	AACAGGTACAGCAAAATCC	TAAAGTCACTTTTAAGATG	TCAGAAATTAATGATGCTGA	60
Db	1164	AACAGGTACAGCAAAATCC	TAAAGTCACTTTTAAGATG	TCAGAAATTAATGATGCTGA	1223
Qy	61	AGGAAGATCATAGATGTC	ATAGTAAGGAACCTTAAT	TCAACCTTTTGAGATCTATT	120
Db	1224	AGGAAGATCATAGATGTC	ATAGTAAGGAACCTTAAT	TCAACCTTTTGAGATCTATT	1283

Qy	121	AGAGTTGATATATATCCAGAGCTGGATGGATCTCTCGAGGAAATATGCTTGGTCCAT	180
Db	1284	AGAGTTGATATATATCCAGAGCTGGATGGATCTCTCGAGGAAATATGCTTGGTCCAT	1343
Qy	181	CCTACTAGATCGCTCCAGACTCGGCTACAGATAGTGTGATCTCACTGAAATATTTAT	240
Db	1344	CCTACTAGATCGCTCCAGACTCGGCTACAGATAGTGTGATCTCACTGAAATATTTAT	1403
Qy	241	CCAGTAGAAGATGATGTTATGGAAAGGAGAGACTCAATGAGTCAGTGGCTGATCTGT	300
Db	1404	CCAGTAGAAGATGATGTTATGGAAAGGAGAGACTCAATGAGTCAGTGGCTGATCTGT	1463
Qy	301	GAGCCCACTAAATATCTATGAAGAAACAAACAGACATCTGGAATAATATCCATGACATCTT	360
Db	1464	GAGCCCACTAAATATCTATGAAGAAACAAACAGACATCTGGAATAATATCCATGACATCTT	1523
Qy	361	TCATGTTTTCCCAAGTCAAGAGAGGAAATGAGTTATTTTTTGCCTCTGAATGCAA	420
Db	1524	TCATGTTTTCCCAAGTCAAGAGAGGAAATGAGTTATTTTTTGCCTCTGAATGCAA	1583
Qy	421	AACAGGTTCCGTCATTTATACAAATATACATCTATTTTAAAGGAAAGCAATATAAAG	480
Db	1584	AACAGGTTCCGTCATTTATACAAATATACATCTATTTTAAAGGAAAGCAATATAAAG	1643
Qy	481	ATCCAGTGGGCTGCCCTCCCAAGTGATTTCAAGTGCTCTATCAAGAGGAGATAGC	540
Db	1644	ATCCAGTGGGCTGCCCTCCCAAGTGATTTCAAGTGCTCTATCAAGAGGAGATAGC	1703
Qy	541	AATTAACAGTGGTGAATGGGAAGTTCTTGGCGGCGATGGATCTAATATCCAGTTGATGA	600
Db	1704	AATTAACAGTGGTGAATGGGAAGTTCTTGGCGGCGATGGATCTAATATCCAGTTGATGA	1763
Qy	601	AGTCAGAGGCTGGTATATTTTGAAGGCAACAAAGACTCCCTTTTAGAGCATCACCTGTA	660
Db	1764	AGTCAGAGGCTGGTATATTTTGAAGGCAACAAAGACTCCCTTTTAGAGCATCACCTGTA	1823
Qy	661	CGTAGTCAGTTACGTAAATCTCGAGAGGCTGACAAGGCTGACTACCGTGGCTACTCACA	720
Db	1824	CGTAGTCAGTTACGTAAATCTCGAGAGGCTGACAAGGCTGACTACCGTGGCTACTCACA	1883
Qy	721	TTCTTGCTGCATCAGTCAGCACTGTGACTTCTTTTAAAGTAAATAGTAAACCAAGAGAA	780
Db	1884	TTCTTGCTGCATCAGTCAGCACTGTGACTTCTTTTAAAGTAAATAGTAAACCAAGAGAA	1943
Qy	781	TCACACACTGTGTCCCTTTTACAAGCTATCAAGTCTCGAAGATGACCCAACTTGCAAAAC	840
Db	1944	TCACACACTGTGTCCCTTTTACAAGCTATCAAGTCTCGAAGATGACCCAACTTGCAAAAC	2003
Qy	841	AAAGGAATTTTGGGCCACCAATTTTGGATTTCAGCAGGTCCTCTTCTGACTATACTCCTCC	900
Db	2004	AAAGGAATTTTGGGCCACCAATTTTGGATTTCAGCAGGTCCTCTTCTGACTATACTCCTCC	2063
Qy	901	AGAAATTTTCTCTTTTGAAGTACTACTGGATTTACATTTGATGGGATGCTCTACAAGCC	960
Db	2064	AGAAATTTTCTCTTTTGAAGTACTACTGGATTTACATTTGATGGGATGCTCTACAAGCC	2123
Qy	961	TCATGATCTACAGCTGGGAAGAAATATCCTACTGTGCTGTTCATATATGTTGCTCTCA	1020
Db	2124	TCATGATCTACAGCTGGGAAGAAATATCCTACTGTGCTGTTCATATATGTTGCTCTCA	2183
Qy	1021	-----	1020
Db	2184	GGTGCAAGTTGGTGAATAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCCTAGC	2243
Qy	1021	-----	1020
Db	2244	CTCTCTAGGTTATGTGGTTGTAGTAGACAAACAGGGGATTCCTGTACCGAGGGCTTAA	2303
Qy	1021	-----	1020
Db	2304	ATTGGAAGCGCCTTTAAATATAAAATGGGTCAAAATAGAAATTGACGATCAGGTGGAAG	2363
Qy	1021	-----	1020
Db	2364	ACTCCAATATCTAGCTTCTCGATATGATTTTCACTGACTTAGATCGTGGGCAATCCACGG	2423
Qy	1021	-----	1020
Db	2424	CTGGTCCCTATPGGAGGATACCTCTCCTGATGGCATTAATGACAGAGGTGAGATATCTTCAG	2483
Qy	1021	GGTTGCTATTTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGA	1080
Db	2484	GGTTGCTATTTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGA	2543
Qy	1081	ACGTTATATGGGTCAACCCCTGACAGAAATGAAACAGGGCTATTACTTAGGATCTGTGGCCAT	1140
Db	2544	ACGTTATATGGGTCAACCCCTGACAGAAATGAAACAGGGCTATTACTTAGGATCTGTGGCCAT	2603
Qy	1141	GCAAGCAGAAAAAGTTCCCTCTGAAACCAATCGTTTACTGTCTTTACATGCTTCTCTGGA	1200
Db	2604	GCAAGCAGAAAAAGTTCCCTCTGAAACCAATCGTTTACTGTCTTTACATGCTTCTCTGGA	2663
Qy	1201	TGAGAAATGTCATTTTGGCACATACCAAGTATATTTACTGAGTTTCTTGTAGGGCTGGAAA	1260
Db	2664	TGAGAAATGTCATTTTGGCACATACCAAGTATATTTACTGAGTTTCTTGTAGGGCTGGAAA	2723
Qy	1261	GCCATATGATTTACAGATCTATCCTCAGGAGAGACACAGCATTAAGATTCCTGAATCGGG	1320
Db	2724	GCCATATGATTTACAGATCTATCCTCAGGAGAGACACAGCATTAAGATTCCTGAATCGGG	2783
Qy	1321	AGACATTATGATCGATCTCTTTTGGCACTACCTTCAAGAAACCTTGGATCAGCTATTGC	1380
Db	2784	AGACATTATGATCGATCTCTTTTGGCACTACCTTCAAGAAACCTTGGATCAGCTATTGC	2843
Qy	1381	TGCTCTAAAGTGATATAATTTTGACCTGTGTGAGAACTCTCTGGTATACACTGCTATT	1440
Db	2844	TGCTCTAAAGTGATATAATTTTGACCTGTGTGAGAACTCTCTGGTATACACTGCTATT	2903
Qy	1441	AACCAATGAGGAGGTTTAAATCAACAGAAACACAGAAATGATCATCATATTTTGATACC	1500
Db	2904	AACCAATGAGGAGGTTTAAATCAACAGAAACACAGAAATGATCATCATATTTTGATACC	2963
Qy	1501	TGCCATGTAACTACTACTCTCTGAAATATAATGTGGTGCCTGCGAGGGGTCTACGGTTGT	1560
Db	2964	TGCCATGTAACTACTACTCTCTGAAATATAATGTGGTGCCTGCGAGGGGTCTACGGTTGT	3023
Qy	1561	GGTAGTAATCTAATACCTTTAAACCCACATGCTCAAAATCAAAATGATACATATTCCTGAGA	1620
Db	3024	GGTAGTAATCTAATACCTTTAAACCCACATGCTCAAAATCAAAATGATACATATTCCTGAGA	3083
Qy	1621	GACCCAGCAATACCAAGAAATTTACTAAAAAATAAATAAATAAATAAATAAATAAATAA	1657
Db	3084	GACCCAGCAATACCAAGAAATTTACTAAAAAATAAATAAATAAATAAATAAATAAATAA	3120
RESULT 6			
US-09-976-674-22			
; Sequence 22, Application US/09976674			
; Patent No. 6844180			
; GENERAL INFORMATION:			
; APPLICANT: Qi Steve			
; APPLICANT: Akinsanya, Karen			
; APPLICANT: Riviere, Jean-Louis			
; APPLICANT: Junien, Jean-Louis			
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV			
; FILE REFERENCE: 70669			
; CURRENT APPLICATION NUMBER: US/09/976,674			
; PRIOR FILING DATE: 2001-10-12			
; PRIOR APPLICATION NUMBER: US 60/240,117			
; PRIOR FILING DATE: 2000-10-12			
; NUMBER OF SEQ ID NOS: 61			
; SOFTWARE: PatentIn version 3.1			
; SEQ ID NO 22			
; LENGTH: 4685			
; TYPE: DNA			
; ORGANISM: Homo sapiens			

US-09-976-674-22

Query Match 80.5%; Score 1343.6; DB 3; Length 4685;
Best Local Similarity 87.0%; Pred. No. 0;
Matches 1589; Conservative 0; Mismatches 79; Indels 158; Gaps 4;

QY 1 AACAGGTACAGCAATCCCTAAAGTCACCTTTTAAAGATGTCAGAAATTAATGATGATGCTGA 60
DB 1164 AACAGGTACAGCAATCCCTAAAGTCACCTTTTAAAGATGTCAGAAATTAATGATGATGCTGA 1223

QY 61 AGGAGGATCATAGATGTCATAGATAAGGAACCTTAATCAACCTTTTGAGATTCATTTTGA 120
DB 1224 AGGAGGATCATAGATGTCATAGATAAGGAACCTTAATCAACCTTTTGAGATTCATTTTGA 1283

QY 121 AGGAGTTGAATATATTGCGAGAGCTGGATGACCTCTGAGGAAATATGCTTGGTCCAT 180
DB 1284 AGGAGTTGAATATATTGCGAGAGCTGGATGACCTCTGAGGAAATATGCTTGGTCCAT 1343

QY 181 CCTACTAGATCCGCTCCAGACTCGCCTACAGATAGTGTGATCTCACCTGAAATATTAT 240
DB 1344 CCTACTAGATCCGCTCCAGACTCGCCTACAGATAGTGTGATCTCACCTGAAATATTAT 1403

QY 241 CCCAGTAGAAGATGATGTTATGGAAGGACAGACTCATTTGAGTCAGTGCCTGATTCGT 300
DB 1404 CCCAGTAGAAGATGATGTTATGGAAGGACAGACTCATTTGAGTCAGTGCCTGATTCGT 1463

QY 301 GAGGCCACTAATATTCTATGAAGAAACAACAGACATCTGGATAAATATCCATGACATCTT 360
DB 1464 GAGGCCACTAATATTCTATGAAGAAACAACAGACATCTGGATAAATATCCATGACATCTT 1523

QY 361 TCATGTTTTTCCCAAGTTCAGAGAGGAAATTTGAGTTATTTTTTGGCTCTGAATGCAA 420
DB 1524 TCATGTTTTTCCCAAGTTCAGAGAGGAAATTTGAGTTATTTTTTGGCTCTGAATGCAA 1583

QY 421 AACAGTTTCCGTCTATTTATACAAATTTACATCTATTTTAAAGAAAGCAATATATAACG 480
DB 1584 AACAGTTTCCGTCTATTTATACAAATTTACATCTATTTTAAAGAAAGCAATATATAACG 1643

QY 481 ATCCAGTGTGGGCTGCCTGCTCCCAAGTGATTTCAAGTGCTCTATCAAGAGAGATAGC 540
DB 1644 ATCCAGTGTGGGCTGCCTGCTCCCAAGTGATTTCAAGTGCTCTATCAAGAGAGATAGC 1703

QY 541 AATTACAGTGTGATGGAAGTTCTTTGGCGGCATGGATCTAATATCAAGTTGATGA 600
DB 1704 AATTACAGTGTGATGGAAGTTCTTTGGCGGCATGGATCTAATATCAAGTTGATGA 1763

QY 601 AGTCAGAGGCTGGTATATTTTGAAGGACCAAAAGACTCCCTTTTAGAGCATCACCTGTA 660
DB 1764 AGTCAGAGGCTGGTATATTTTGAAGGACCAAAAGACTCCCTTTTAGAGCATCACCTGTA 1823

QY 661 CGTAGTCAGTTACGTAAATCCCTGGAGAGGTGACAGGCTGACCTGACCGTGGCTACTACA 720
DB 1824 CGTAGTCAGTTACGTAAATCCCTGGAGAGGTGACAGGCTGACCTGACCGTGGCTACTACA 1883

QY 721 TTCTTCTGTCATCAGTCAGCACTGTGACTTCTTTTAAAGTAAGTATAGTAAACAGAGAA 780
DB 1884 TTCTTCTGTCATCAGTCAGCACTGTGACTTCTTTTAAAGTAAGTATAGTAAACAGAGAA 1943

QY 781 TCCACACTGTGTCCCTTTTACAGCTATCAAGTCTGAAGATGACCCAACTTGGAAGAAC 840
DB 1944 TCCACACTGTGTCCCTTTTACAGCTATCAAGTCTGAAGATGACCCAACTTGGAAGAAC 2003

QY 841 AAAGGAATTTGGGCCACCAATTTGGATTCTAG-----CAGGTCTCTTCCCTGACTACTC 896
DB 2004 AAAGGAATTTGGGCCACCAATTTGGATTCTAGCTCCTCAGGTGCAAGTTGGTGAATAATCGG 2063

QY 897 CTCCA-----GAAATTTTCTCTTTTGAAGTACTACTGGAATTTTACATTTGATGGGATGC 950
DB 2064 TTTTAAAGGAGTCAAGTATTTCCGCTTGATACCTTAGCCTCTTAGGTTATGTTGTA 2123

QY 951 TCTACAGGCTCATGATCTACAG----- 973
DB 2124 GTGATAGACAACAGGGGATCTCTGTCAACGAGGCTTAAATTTGAAGGCGCTTTTAAATAT 2183

QY 974 ----- 973
DB 2184 AAAATGGGTCAAATAGAAATTCACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGA 2243

QY 974 -----CCTGGAAAGAAATATCTTACTGTGCTGTTTATATATATATGTTGGTCTC 1019
DB 2244 TATGATTTTATTGACTTAGATCTGTGGGCATCCACGGCTGGTCTATATGGAGGATACCTC 2303

QY 1020 -----AGTTTGTCTTATTTGCTGGGGCCCCA 1042
DB 2304 TCCTGATGCGCATTAATGACGAGGTACAGATATCTTCAGGGTTGCTATTTGCTGGGGCCCCA 2363

QY 1043 GTCACCTCTGTGGATCTTCTATGATACAGGATACACGGAAACGTTTATATGGGTACCCCTGAC 1102
DB 2364 GTCACCTCTGTGGATCTTCTATGATACAGGATACACGGAAACGTTTATATGGGTACCCCTGAC 2423

QY 1103 CAGAATGAACAGGGCTATTACTTAGATCTGTGGCATGCAAGCAGAGAAAGTTCCCTCTCT 1162
DB 2424 CAGAATGAACAGGGCTATTACTTAGGATCTGTGGCATGCAAGCAGAGAAAGTTCCCTCTCT 2483

QY 1163 GAACCAAAATCGTTTACTGCTCTTACATGTTTCTTGGATGAGAATGTCATTTTGCACAT 1222
DB 2484 GAACCAAAATCGTTTACTGCTCTTACATGTTTCTTGGATGAGAATGTCATTTTGCACAT 2543

QY 1223 ACCAGTATATTACTGAGTTTTTTAGTGAGGCTGGAAAGCCATATGATTTTACAGATCTAT 1282
DB 2544 ACCAGTATATTACTGAGTTTTTTAGTGAGGCTGGAAAGCCATATGATTTTACAGATCTAT 2603

QY 1283 CCTCAGAGAGACACAGCATTAAGAGTTCTGATCGGAGAGAACATTTATGAATCTGATCTT 1342
DB 2604 CCTCAGAGAGACACAGCATTAAGAGTTCTGATCGGAGAGAACATTTATGAATCTGATCTT 2663

QY 1343 TTGCACCTACTCTTCAAGAAACCTTGGATCAGCTATTTGCTCTTAAAGTGATATAATTT 1402
DB 2664 TTGCACCTACTCTTCAAGAAACCTTGGATCAGCTATTTGCTCTTAAAGTGATATAATTT 2723

QY 1403 TGACCTGTGTAGAACTCTCTGGTATACACTGGCTATTTTAAACCAATGAGGAGGTTTAAATC 1462
DB 2724 TGACCTGTGTAGAACTCTCTGGTATACACTGGCTATTTTAAACCAATGAGGAGGTTTAAATC 2783

QY 1463 AACAGAAACACAGAAATGATCATCATTTTGATACCTGCGCATGTAACATCTACTCTG 1522
DB 2784 AACAGAAACACAGAAATGATCATCATTTTGATACCTGCGCATGTAACATCTACTCTG 2843

QY 1523 AAAATAAATGCTGTGCGCATGTCAGGGGTCTACGGTTTGTGGTAGTAATCTAATACTTAAAC 1582
DB 2844 AAAATAAATGCTGTGCGCATGTCAGGGGTCTACGGTTTGTGGTAGTAATCTAATACTTAAAC 2903

QY 1583 CCCACATGCTCAAAATCAAAATGATACATATTTCTGAGAGACCCAGCAATACCAATAAGAT 1642
DB 2904 CCCACATGCTCAAAATCAAAATGATACATATTTCTGAGAGACCCAGCAATACCAATAAGAT 2963

QY 1643 TACTAAAAAATAAAAAAAAAAAAAA 1668
DB 2964 TACTAAAAAATAAAAAAAAAAAAAA 2989

RESULT 7
US-09-976-594-1103
; Sequence 1103, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
; APPLICANT: Furness, Michael
; APPLICANT: Buchbinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/09/976,594
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,403
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program

```
; SEQ ID NO 1103
; LENGTH: 2797
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Inceyte ID No. 6673549 977951.1
US-09-976-594-1103

Query Match      80.1%; Score 1337; DB 3; Length 2797;
Best Local Similarity 84.5%; Pred. No. 0;
Matches 1661; Conservative 0; Mismatches 0; Indels 304; Gaps 2;

Qy      1 AACAGGTACAGCAAAATCTTAAAGTCACTTTTAAAGATGTCAGAAAATAATGATGCTGA 60
Db      |||
Qy      831 AACAGGTACAGCAAAATCTTAAAGTCACTTTTAAAGATGTCAGAAAATAATGATGCTGA 890
Db      |||
Qy      61 AGAAGGATCATAGATGTCATAGATAAGGAACCTAAATCAACCTTTTGGAGATTCTATTGGA 120
Db      |||
Qy      891 AGAAGGATCATAGATGTCATAGATAAGGAACCTAAATCAACCTTTTGGAGATTCTATTGGA 950
Db      |||
Qy      121 AGGAGTTGAATATATTGCCAGAGCTGGATGGACTCCTGAGGGAATAA-----TGCTTGGT 176
Db      |||
Qy      951 AGGAGTTGAATATATTGCCAGAGCTGGATGGACTCCTGAGGGAATAAAGTGTCTTGGT 1010
Db      |||
Qy      177 CCATCTCTAGATGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAATTAT 236
Db      |||
Qy      1011 CCATCTCTAGATGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAATTAT 1070
Db      |||
Qy      237 TTATCCCACTAGAGATGATGTTATGGAAGGCGAGAGACTCATTGAGTCACTGCTGATT 296
Db      |||
Qy      1071 TTATCCCACTAGAGATGATGTTATGGAAGGCGAGAGACTCATTGAGTCACTGCTGATT 1130
Db      |||
Qy      297 CTGTGAGCGCACTAATATTATCTATGAAAGAACACAGACATCTGGATTAATATCCATGACA 356
Db      |||
Qy      1131 CTGTGAGCGCACTAATATTATCTATGAAAGAACACAGACATCTGGATTAATATCCATGACA 1190
Db      |||
Qy      357 TCTTTCATGTTTTTCCCAAGTCAAGAGGAAATGAGTTATTTTGGCTCTGAAT 416
Db      |||
Qy      1191 TCTTTCATGTTTTTCCCAAGTCAAGAGGAAATGAGTTATTTTGGCTCTGAAT 1250
Db      |||
Qy      417 GCAAAACAGGTTTCCGTCATTATACAAAATTAATCATCTATTTTAAAGGAAAGCAAAATATA 476
Db      |||
Qy      1251 GCAAAACAGGTTTCCGTCATTATACAAAATTAATCATCTATTTTAAAGGAAAGCAAAATATA 1310
Db      |||
Qy      477 AACGATPCCAGTGGTGGCTGCTGCCAAGTGATTTCAAGTGTCTCTATCAAGAGGAGA 536
Db      |||
Qy      1311 AACGATPCCAGTGGTGGCTGCTGCCAAGTGATTTCAAGTGTCTCTATCAAGAGGAGA 1370
Db      |||
Qy      537 TAGCAATTACCAAGTGGTGAATGGGAAGTTCTTGGCCGGCATGGATCTAATATCCAAAGTTG 596
Db      |||
Qy      1371 TAGCAATTACCAAGTGGTGAATGGGAAGTTCTTGGCCGGCATGGATCTAATATCCAAAGTTG 1430
Db      |||
Qy      597 ATGAAGTCAGAGGCTGGTATATTTTGAAGGCACCAAGACTCCCTTTAGAGCATCACC 656
Db      |||
Qy      1431 ATGAAGTCAGAGGCTGGTATATTTTGAAGGCACCAAGACTCCCTTTAGAGCATCACC 1490
Db      |||
Qy      657 TGTACGTAGTCAGTTACGTAATCTCGAGAGGTTGACAAGCTGACCTGACCTGGCTACT 716
Db      |||
Qy      1491 TGTACGTAGTCAGTTACGTAATCTCGAGAGGTTGACAAGCTGACCTGACCTGGCTACT 1550
Db      |||
Qy      717 GACATTCCTGTCGATCAGTCAGACCTGTCCTTTTAAAGTAAGTATAGTAACCAAGA 776
Db      |||
Qy      1551 GACATTCCTGTCGATCAGTCAGACCTGTCCTTTTAAAGTAAGTATAGTAACCAAGA 1610
Db      |||
Qy      777 AGAATCCCACTGTGTGCTCTTTTACAGCTATCAAGTCTGGAAGATGCCAACTTGA 836
Db      |||
Qy      1611 AGAATCCCACTGTGTGCTCTTTTACAGCTATCAAGTCTGGAAGATGCCAACTTGA 1670
Db      |||
Qy      837 AAACAAAGGAATTTTGGGCCCAACCAATTTTGGATTACAGAGGTCCTCTTCTGCTACTATCTC 896
Db      |||
Qy      1671 AAACAAAGGAATTTTGGGCCCAACCAATTTTGGATTACAGAGGTCCTCTTCTGCTACTATCTC 1730
Db      |||
```

```
Qy      897 CTCAGGAATTTTCTCTTTTGAAGTACTACTGGATTTTACATTTGATGGGATGCTCTACA 956
Db      |||
Qy      1731 CTCAGGAATTTTCTCTTTTGAAGTACTACTGGATTTTACATTTGATGGGATGCTCTACA 1790
Db      |||
Qy      957 AGCCTCATGATCTACAGCCTGGAAAGAAATATCTTACTGTGCTGTTTATATATATGTTGCTC 1016
Db      |||
Qy      1791 AGCCTCATGATCTACAGCCTGGAAAGAAATATCTTACTGTGCTGTTTATATATATGTTGCTC 1850
Db      |||
Qy      1017 CTC----- 1020
Db      |||
Qy      1851 CTCAGGTGCAGTTGGTGAATAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCC 1910
Db      |||
Qy      1021 ----- 1020
Db      |||
Qy      1911 TAGCCTCTCTAGGTTATGTGTTAGTAGATAGACAAACAGGGGATCTCTGTCACCGAGGC 1970
Db      |||
Qy      1021 ----- 1020
Db      |||
Qy      1971 TTAAATTTGAAGGCGCTTTTAAATATATAAAATGCGGTCAAATAGAAATTTGACGATCAGGTGG 2030
Db      |||
Qy      1021 ----- 1020
Db      |||
Qy      2031 AAGGACTCCAATATCTAGCTTCTCGATATGATTTTCATTGACTTAGATCGTGTGGGCATCC 2090
Db      |||
Qy      1021 ----- 1020
Db      |||
Qy      2091 ACGGCTGGTCTTATGGAGGATACCTCTCCGTGATGGCATTAATGCAGAGGTCAGATATCT 2150
Db      |||
Qy      1021 ----- 1020
Db      |||
Qy      2151 TCAGGTTGTTGATGCTGTGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACA 2210
Db      |||
Qy      1077 CGGAACTGTTATGAGGTCACTGACCAAGATGAACAGGGCTATTTACTTTAGGATCTCTGG 1136
Db      |||
Qy      2211 CGGAACTGTTATGAGGTCACTGACCAAGATGAACAGGGCTATTTACTTTAGGATCTCTGG 2270
Db      |||
Qy      1137 CCATGCAAGCAGAAAAGTTTCCCTCTGAAACCAAAATCGTTTACTGCTCTTACATGTTTCC 1196
Db      |||
Qy      2271 CCATGCAAGCAGAAAAGTTTCCCTCTGAAACCAAAATCGTTTACTGCTCTTACATGTTTCC 2330
Db      |||
Qy      1197 TGGATGAGAAATGCTCCATTTTGACATACCACTAGTATATTTACTGAGTTTTTGTAGTGGGCTG 1256
Db      |||
Qy      2331 TGGATGAGAAATGCTCCATTTTGACATACCACTAGTATATTTACTGAGTTTTTGTAGTGGGCTG 2390
Db      |||
Qy      1257 GAAAGCCATATGATTTACAGATCTATCTCAGGAGAGACACAGCATTAAGAGTTTCTCGAAT 1316
Db      |||
Qy      2391 GAAAGCCATATGATTTACAGATCTATCTCAGGAGAGACACAGCATTAAGAGTTTCTCGAAT 2450
Db      |||
Qy      1317 CGGAGAACTATGAACTGCACTTTTGGCACTACCTTCAAGAAACCTTGGATCAGGTA 1376
Db      |||
Qy      2451 CGGAGAACTATGAACTGCACTTTTGGCACTACCTTCAAGAAACCTTGGATCAGGTA 2510
Db      |||
Qy      1377 TTGCTGCTCTAAAGTGATATAATTTTGACCTGTGTAGAACTCTCTCGTATACACTGGCT 1436
Db      |||
Qy      2511 TTGCTGCTCTAAAGTGATATAATTTTGACCTGTGTAGAACTCTCTCGTATACACTGGCT 2570
Db      |||
Qy      1437 ATTTAACCAAAATGAGAGGTTTAAATCAAGAAACACAGAAATTTGATCATCATTTTGA 1496
Db      |||
Qy      2571 ATTTAACCAAAATGAGAGGTTTAAATCAAGAAACACAGAAATTTGATCATCATTTTGA 2630
Db      |||
Qy      1497 TACTGCTCAATGTAACATCTACTCTGAAAATAAATGTTGGTGCCCATGAGGGGCTACGGT 1556
Db      |||
Qy      2631 TACTGCTCAATGTAACATCTACTCTGAAAATAAATGTTGGTGCCCATGAGGGGCTACGGT 2690
Db      |||
Qy      1557 TTGTGTAGTAAATCTAATACCTTAAACCCACATGCTCAAAATCAAAATGATACATATTCCT 1616
Db      |||
Qy      2691 TTGTGTAGTAAATCTAATACCTTAAACCCACATGCTCAAAATCAAAATGATACATATTCCT 2750
Db      |||
Qy      1617 GAGAGACCCAGCAATACCATTAAGAAATTTACTAAAAAATAAATAAATAAATAAATAAATAA 1661
Db      |||
Qy      2751 GAGAGACCCAGCAATACCATTAAGAAATTTACTAAAAAATAAATAAATAAATAAATAAATAA 2795
Db      |||
```

RESULT 8	
US-09-976-674-2 ;	
; Sequence 2, Application US/09976674	
; Patent No. 6844180	
; GENERAL INFORMATION:	
; APPLICANT: Qi, Steve	
; APPLICANT: Akinsanya, Karen	
; APPLICANT: Riviere, Pierre	
; APPLICANT: Junien, Jean-Louis	
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV	
; FILE REFERENCE: 70669	
; CURRENT APPLICATION NUMBER: US/09/976,674	
; CURRENT FILING DATE: 2001-10-12	
; PRIOR APPLICATION NUMBER: US 60/240,117	
; PRIOR FILING DATE: 2000-10-12	
; NUMBER OF SEQ ID NOS: 61	
; SOFTWARE: PatentIn version 3.1	
; SEQ ID NO 2	
; LENGTH: 2671	
; TYPE: DNA	
; ORGANISM: Homo sapiens	
US-09-976-674-2	
Query Match 65.1%; Score 1085.8; DB 3; Length 2671;	
Best Local Similarity 82.2%; Pred. No. 7.le-289;	
Matches 1397; Conservative 0; Mismatches 2; Indels 300; Gaps 1;	
QY	1 AACAGGTACAGCAAAATCCTAAAGTCACCTTTTAAAGATGTCAGAAATAATGATTGATGCTGA 60
DB	958 AACAGGTACAGCAAAATCCTAAAGTCACCTTTTAAAGATGTCAGAAATAATGATTGATGCTGA 1017
QY	61 AGGAAGGATCATAGATGTCATAGATAAGGAACATAATCAACCTTTTGAGATTCATTGTA 120
DB	1018 AGGAAGGATCATAGATGTCATAGATAAGGAACATAATCAACCTTTTGAGATTCATTGTA 1077
QY	121 AGGAGTTGAATATATTGCCAGAGCTGGATGGATCTCTCAGGAGAAATATGCTTGGTCCAT 180
DB	1078 AGGAGTTGAATATATTGCCAGAGCTGGATGGATCTCTCAGGAGAAATATGCTTGGTCCAT 1137
QY	181 CCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACTGAATATTATTAT 240
DB	1138 CCTACTAGATCGCTCCAGACTCGCTCGAGATAGTGTGATCTCACTGAATATTATTAT 1197
QY	241 CCCAGTAGAAGATGATGTTTATGGAAGGCAGAGACTCAATTGAGTCAGTGCCTGATTCGT 300
DB	1198 CCCAGTAGAAGATGATGTTTATGGAAGGCAGAGACTCAATTGAGTCAGTGCCTGATTCGT 1257
QY	301 GAGCCACATAATTATCTATGAAAGAAACACAGACATCTGGATAAATATCCATGACATCTT 360
DB	1258 GAGCCACATAATTATCTATGAAAGAAACACAGACATCTGGATAAATATCCATGACATCTT 1317
QY	361 TCATGTTTTTCCCAAGTCACGAGAGAAATTCAGTTTATTTTTTGCCTCTGAATGCAA 420
DB	1318 TCATGTTTTTCCCAAGTCACGAGAGAAATTCAGTTTATTTTTTGCCTCTGAATGCAA 1377
QY	421 AACAGGTTTCCTGTCATTTATACAAAATTTACATCTATTTTTTAAAGGAAAGCAAAATATAACG 480
DB	1378 AACAGGTTTCCTGTCATTTATACAAAATTTACATCTATTTTTTAAAGGAAAGCAAAATATAACG 1437
QY	481 ATCCAGTGGGGCTGCCTGCTCCAAAGTGATTTTCAAGTGTCTTATCAAGAGGAGATAGC 540
DB	1438 ATCCAGTGGGGCTGCCTGCTCCAAAGTGATTTTCAAGTGTCTTATCAAGAGGAGATAGC 1497
QY	541 AATTACAGTGGTGAATGGGAAGTTCTTGGCCGGATGGATCTAATATCCAAGTTGATGA 600
DB	1498 AATTACAGTGGTGAATGGGAAGTTCTTGGCCGGATGGATCTAATATCCAAGTTGATGA 1557
QY	601 AGTCAGAGGCTGGTATATTTTGAAGGACCAAGACTCCCTTTTAGAGCATCACCTGTGA 660
DB	1558 AGTCAGAGGCTGGTATATTTTGAAGGACCAAGACTCCCTTTTAGAGCATCACCTGTGA 1617
QY	661 CGTAGTCAGTTACGTAATTCCTGGAGAGTGACAAGGCTGACTGACCGTGGCTACTCACA 720

```
; Sequence 4, Application US/10070464
; Patent No. 6881564
; GENERAL INFORMATION:
; APPLICANT: ABBOTT, Catherine Anne
; APPLICANT: CORRELL, Mark Douglas
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES
; FILE REFERENCE: GH-007
; CURRENT APPLICATION NUMBER: US/10/070,464
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: PCT/AU00/01085
; PRIOR FILING DATE: 2000-09-11
; PRIOR APPLICATION NUMBER: AU PQ5709
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: AU PQ2762
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1197
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-10-070-464-4

Query Match      52.8%; Score 881; DB 3; Length 1197;
Best Local Similarity 87.2%; Pred. No. 1.4e-232;
Matches 1044; Conservative 0; Mismatches 0; Indels 153; Gaps 1;

QY 618 ATTTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTAGCTAGTCAGTTACGTAA 677
DB 1 ATTTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTAGCTAGTCAGTTACGTAA 60
QY 678 ATCTCGAGAGGTGACCAAGCTGACTGACCGTGGTACTACATCTTGTGTCATGATG 737
DB 61 ATCTCGAGAGGTGACCAAGCTGACTGACCGTGGTACTACATCTTGTGTCATGATG 420
QY 738 AGCACTGTGACTCTTTTAAAGTATAGTAACTACCAAGAAATCCACATGTGTGCC 797
DB 121 AGCACTGTGACTCTTTTAAAGTATAGTAACTACCAAGAAATCCACATGTGTGCC 180
QY 798 TTTTCAAGCTATCAAGTCTGAAGATGACCAACTTGCACCAAGAAATTTTGGGCCA 857
DB 181 TTTTCAAGCTATCAAGTCTGAAGATGACCAACTTGCACCAAGAAATTTTGGGCCA 240
QY 858 CCATTTTGGATTACAGAGTCTCTCTCTGACTATATCTCTCCAGAAATTTCTCTTTTG 917
DB 241 CCATTTTGGATTACAGAGTCTCTCTCTGACTATATCTCTCCAGAAATTTCTCTTTTG 300
QY 918 AAAGTACTACTGATTACATTTGATGGATGCTCTACAAGCCTCATGATCTACAGCTG 977
DB 301 AAAGTACTACTGATTACATTTGATGGATGCTCTACAAGCCTCATGATCTACAGCTG 360
QY 978 GAAAGAAATATCTACTGTGCTGTTTCATATATATGTTGGTCTCTCAGGGTCAAAATG 1020
DB 361 GAAAGAAATATCTACTGTGCTGTTTCATATATATGTTGGTCTCTCAGGGTCAAAATG 420
QY 1021 ----- 1020
DB 421 ACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTTCATTGACTGATG 480
QY 1021 ----- 1020
DB 481 GTGTGGGCATCCAGCGTGGTCTATGAGATACCTCTCTCGTGAAGGCATTAATCCAGA 540
QY 1021 ----- 1020
DB 541 GGTGAGATATCTTCAAGGTTGCTATGCTGGGGGCCAGTCACTCTGTGGATCTTCTATG 1064
QY 1065 ATACAGGATACACCGAACTTATATGGGTCACTGACCTGACCAAGAAATGAACAGGGCTATT 1124
DB 601 ATACAGGATACACCGAACTTATATGGGTCACTGACCTGACCAAGAAATGAACAGGGCTATT 660
QY 1125 TAGGATCTGTGGCCATGCAAGCAAGAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCT 1184
```

```
DB 661 TAGGATCTGTGGCCATGCAAGCAAGAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCT 720
QY 1185 TACATGTTTCTCTGGATGAGATGTCATTTTGCACATACCAAGTATATATTACTGAGTTTTT 1244
DB 721 TACATGTTTCTCTGGATGAGATGTCATTTTGCACATACCAAGTATATATTACTGAGTTTTT 780
QY 1245 TAGTGAGGCTGGAAGGCCATATGATTTTACAGATCTATCTCTCAGGAGAGACACAGCATAA 1304
DB 781 TAGTGAGGCTGGAAGGCCATATGATTTTACAGATCTATCTCTCAGGAGAGACACAGCATAA 840
QY 1305 GAGTTCTCTGAATCGGAGAGAACATTATGAATCTCATCTTTTGACACTACCTTCAAGAAACC 1364
DB 841 GAGTTCTCTGAATCGGAGAGAACATTATGAATCTCATCTTTTGACACTACCTTCAAGAAACC 900
QY 1365 TTGATCACGTATTTGCTGCTCTAAAGATGATATAATTTTGACCTGTGAGAACTCTCTGG 1424
DB 901 TTGATCACGTATTTGCTGCTCTAAAGATGATATAATTTTGACCTGTGAGAACTCTCTGG 960
QY 1425 TATACACTGGCTATTTTAAACCAATGAGGAGTTTAAATCAACAGAAACACACAGAAATGATC 1484
DB 961 TATACACTGGCTATTTTAAACCAATGAGGAGTTTAAATCAACAGAAACACACAGAAATGATC 1020
QY 1485 ATCACATTTTGTATACCTGCTGATGTAACATCTACTCTGAAATTAATGTTGTCATGCA 1544
DB 1021 ATCACATTTTGTATACCTGCTGATGTAACATCTACTCTGAAATTAATGTTGTCATGCA 1080
QY 1545 GGGGTCTACGTTTGTGCTAGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATG 1604
DB 1081 GGGGTCTACGTTTGTGCTAGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATG 1140
QY 1605 ATCATATTTCTGAGAGACCCAGCAATACCATTAAGAAATTAATTAATTAATTAATTAATTAAT 1661
DB 1141 ATCATATTTCTGAGAGACCCAGCAATACCATTAAGAAATTAATTAATTAATTAATTAATTAAT 1197

RESULT 10
US-09-976-674-14
; Sequence 14, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
; LENGTH: 4309
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-976-674-14

Query Match      45.8%; Score 765.2; DB 3; Length 4309;
Best Local Similarity 77.7%; Pred. No. 2e-200;
Matches 1082; Conservative 0; Mismatches 8; Indels 302; Gaps 1;

QY 579 GATCTAATATCCAGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGGCACCAAGACT 638
DB 1222 GAAGGAAGATCCCAAGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGGCACCAAGACT 1281
QY 639 CCCCTTTAGAGCATCACCTGTACGTAGTCAGTTACGTAAATCTCTGGAGAGTGACAAAGC 698
DB 1282 CCCCTTTAGAGCATCACCTGTACGTAGTCAGTTACGTAAATCTCTGGAGAGTGACAAAGC 1341
QY 699 TGACTACCGTGGCTACTCACATTTCTGCTGCATCAGTCAGCTGTGACTTCTTTTATAA 758
DB 1342 TGACTACCGTGGCTACTCACATTTCTGCTGCATCAGTCAGCTGTGACTTCTTTTATAA 1401
```

QY 759 GTAAAGTATAGTAACAGAGAATCCACACTGTGTGTCCTTTTCAAGCTATCAAGTCCTG 818
Db 1402 GTAAGTATAGTAACAGAGAATCCACACTGTGTGTCCTTTTCAAGCTATCAAGTCCTG 1461
QY 819 AAGATGACCAACTTTGCAAAAACAAGGAATTTTGGGCCACCATTTTGGATTTCAGCAGGTC 878
Db 1462 AAGATGACCAACTTTGCAAAAACAAGGAATTTTGGGCCACCATTTTGGATTTCAGCAGGTC 1521
QY 879 CTCTTCCTGACTATATCTCTCAGAAAATTTCTCTTTTGAAGTACTACTGGATTATACAT 938
Db 1522 CTCTTCCTGACTATATCTCTCAGAAAATTTCTCTTTTGAAGTACTACTGGATTATACAT 1581
QY 939 TGTATGGGATGCTCTACAGCCTCATGATCTACAGCCTGGAGAAATATCTACTGTGC 998
Db 1582 TGTATGGGATGCTCTACAGCCTCATGATCTACAGCCTGGAGAAATATCTACTGTGC 1641
QY 999 TGTTCATATATGTGTGTCCTC----- 1019
Db 1642 TGTTCATATATGTGTGTCCTCAGGTGCGAGTTGGTGAATATCGGTTTAAAGAGTCAA 1701
QY 1020 ----- 1019
Db 1702 GTATTTCCGTTGAATACCTAGCCTCTCTAGGTTATGTGTTAGTGATAGACAACAG 1761
QY 1020 ----- 1019
Db 1762 GGGATCCTGTACCGAGGCTTAAATTTGAAGGCGCTTTAAATATAAAATGGTCAAAT 1821
QY 1020 ----- 1019
Db 1822 AGAAATTGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTCAATGA 1881
QY 1020 ----- 1019
Db 1882 CTTAGATCGTGTGGGCATCCACGGCTGCTCTATGGAGATACCTCTCCCTGATGGCAAT 1941
QY 1020 ----- AGTTGCTATTGCTGGGCCCCAGTCACTCTGTGGAT 1056
Db 1942 AATGCAAGGTCAATATCTCAGGGTTGCTATTGCTGGGCCCCAGTCACTCTGTGGAT 2001
QY 1057 CTTCTATGATACAGGATACCGGAACGTTATATGGGTCAACCTGACAGAAATGAACAGGG 1116
Db 2002 CTTCTATGATACAGGATACCGGAACGTTATATGGGTCAACCTGACAGAAATGAACAGGG 2061
QY 1117 CTAATTAAGTATGTCGCCATGCAAGCAAGAAAGTTCCCTCTGAAACAAATCGTTT 1176
Db 2062 CTAATTAAGTATGTCGCCATGCAAGCAAGAAAGTTCCCTCTGAAACAAATCGTTT 2121
QY 1177 ACTGCTCTTACATGGTTTCTGGATGAGAAATGTCATTTTGCATATACCATGATATTTACT 1236
Db 2122 ACTGCTCTTACATGGTTTCTGGATGAGAAATGTCATTTTGCATATACCATGATATTTACT 2181
QY 1237 GAGTTTATAGTACAGGCTGGAAGCCATATGATTTTACAGATCTATCTCAGAGAGACA 1296
Db 2182 GAGTTTATAGTACAGGCTGGAAGCCATATGATTTTACAGATCTATCTCAGAGAGACA 2241
QY 1297 CAGCATAAGAGTTTCTGAAATCGGAGAACATTTATGAATGCTATTTTGCACACTACCTTCA 1356
Db 2242 CAGCATAAGAGTTTCTGAAATCGGAGAACATTTATGAATGCTATTTTGCACACTACCTTCA 2301
QY 1357 AGAAAACCTTTGGATACAGTATGTCCTCTAAAGTGATATAAATTTTGAACCTGTGTAGAA 1416
Db 2302 AGAAAACCTTTGGATACAGTATGTCCTCTAAAGTGATATAAATTTTGAACCTGTGTAGAA 2361
QY 1417 CTCTCTGGTATACATGGCTATTTAAACCAATGAGAGGTTTAAATCAACAGAAACACAG 1476
Db 2362 CTCTCTGGTATACATGGCTATTTAAACCAATGAGAGGTTTAAATCAACAGAAACACAG 2421
QY 1477 AATTGATCATCATTTTGTATACCTGCGCATGTAAATCTACTCTCTGAAATTAATGTGGT 1536
Db 2422 AATTGATCATCATTTTGTATACCTGCGCATGTAAATCTACTCTCTGAAATTAATGTGGT 2481

QY 1537 GCCATGACGGGCTTACGGTTTGTGGTAGTAATCTTAATACCCCAATGCTCAAA 1596
Db 2482 GCCATGACGGGCTTACGGTTTGTGGTAGTAATCTTAATACCCCAATGCTCAAA 2541
QY 1597 ATCAATGATACATATTTCTGAGAGACCCAGCAATACCATTAAGATTTACTTAAAAA 1656
Db 2542 ATCAATGATACATATTTCTGAGAGACCCAGCAATACCATTAAGATTTACTTAAAAA 2601
QY 1657 AAAAAA 1668
Db 2602 AAAAAA 2613
RESULT 11
US-09-280-116-171
; Sequence 171, Application US/09280116A
; Patent No. 6331427
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: Nucleic Acid Molecules Encoding Human Protease Homologs
; FILE REFERENCE: 5800-24, 035800/176965
; CURRENT APPLICATION NUMBER: US/09/280,116A
; CURRENT FILING DATE: 1999-03-26
; NUMBER OF SEQ ID NOS: 268
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 171
; LENGTH: 823
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: prolyl oligopeptidases
US-09-280-116-171
Query Match 29.6%; Score 493.4; DB 3; Length 823;
Best Local Similarity 93.1%; Pred. No. 8.1e-126;
Matches 625; Conservative 0; Mismatches 11; Indels 35; Gaps 9;
QY 1021 GGTGTCTATTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGA 1080
Db 153 GGTGTCTATTGCTGGGTCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGA 212
QY 1081 ACGTTATATGGGTCAACCTGACAGAAATGAACAGGGCTATTACTTAGGATCTGTGGCCAT 1140
Db 213 ACGTTATATGGGTCAACCTGACAGAAATGAACAGGGCTATTACTTAGGATCTGTGGCCAT 272
QY 1141 GCAAGCAGAAAAGTTCCCTCTGAACCAATCGTTTACTGCTCTTACATGCTTCTCTGGA 1200
Db 273 GCAAGCAGAAAAGTTCCCTCTGAACCAATCGTTTACTGCTCTTACATGCTTCTCTGGA 332
QY 1201 TGGAATGTCCATTTTGCACATACCATGATATTTACTGAGTTTGTAGTGAGGCTGGAAA 1260
Db 333 TGGAATGTCCATTTTGCACATACCATGATATTTACTGAGTTTGTAGTGAGGCTGGAAA 392
QY 1261 GCCATATGATTTAC-----AGATCTATCTCTCA--GGAGAG 1293
Db 393 GCAGTATGATTTACAGTATCTTTATTTTGTGTGTGTAGATCTATCTCAGGAGAG 452
QY 1294 ACA-CAGCATAAGAGTTCC--TGAATCGGAGAAACATTATGAATCTTTT-GCACT 1349
Db 453 ACACCGCATAAGAGTTCCCTGAATCGGGGAGAACATTATGAATCTCTTTTGGCACT 512
QY 1350 ACCTTCAAG-AAAACCTTGGATCA-CGTATTCGTGT-CTAAAAAGTATTAATTTTGCAC 1406
Db 513 ACCTTCAAGAAAAACCTTGGATCACCTTATTCGTGTCTCTAAAAATGATATAATTTTGCAC 572
QY 1407 CTGTGTAGACTCTCT-GGTATACCTGCTATTTTAAACCAATGAGGAGGTTTAAATCAAC 1465
Db 573 CTGTGTAGAACTCTCTGGGTATACACTGGCTATTTAAACCAATGAGGAGGTTTAAATCAAC 632
QY 1466 AGAAAAACAGAAATGATCATCATTTTGTATACCTGCGCATGTAAATCTACTCTCTGAAA 1525
Db 633 AGAAAAACAGAAATGATCATCATTTTGTATACCTGCGCATGTAAATCTACTCTCTGAAA 692

; SEQ ID NO 30		
; LENGTH: 4159		
; TYPE: DNA		
; ORGANISM: Homo sapiens		
us-09-976-674-30		
Query Match		
Best Local Similarity 57.3%; Pred. No. 8.2e-76;		
Matches 587; Conservative 0; Mismatches 435; Indels 3; Gaps 1;		
Qy	2	ACAGGTACAGCAAACTCTAAAGTCACTTTTAAGATGTCAGAAATAATGATGTGCTGAA 61
Db	1288	ACAGGCGACGAAGATCCCAAGATTGCTTTGAACTGGCTGAGTTCAGACTGACAGCCAG 1347
Qy	62	GGAAGGATCATAGATGTCATAGATAAGGAACCTAATTTCAACCTTTTGAGATTCTATTGAA 121
Db	1348	GGCAAGATCTCTCGACCCCGAGAGAGAGCTGGTGCAGCCCTTCAGCTGGCTGTTCCCG 1407
Qy	122	GGAGTTGAATATATTGCCAGAGCTGGATGGACTCTCGAGGAGAAATATGCTTGGTCCATC 181
Db	1408	AAGGTGGAGTACATCGCCAGGGCCGGGTGGACCCGGGATGCAAAATACGGCTGGGCCATG 1467
Qy	182	CTACTAGATCGCTCCAGACTCGCCTACAGATAGTGTGATCTCACCTGAATTATTATC 241
Db	1468	TTCTGGACCGGCCCCAGCAGTGGCTCCAGCTCGTCTCTCCCCCGGCCCTGTTTCATC 1527
Qy	242	CCAGTAGAAGATCATGTTATGGAAGGCGAGAGACTCATTTGAGTCAGTGCCTGATTCTGTG 301
Db	1528	CCGAGCACAGAGATGAGGAGCGGCTAGCCTCTGCCAGAGCTGTCCCCAGGAATGTC 1587
Qy	302	ACGCCACTAATTATCTATGAGAAACACAGACATCTGGATATAATATCCATGACATCTTT 361
Db	1588	CAGCCGTATGTGTGTACAGAGAGTCCCAACGCTGAGTCAATGTTCAATGACATCTTC 1647
Qy	362	CATGTTTTTCCCAAGTCAAGAGAG---GAATTGAGTTTATTTTGGCTCTGATGC 418
Db	1648	TATCCCTTCCCCCAATCAGAGGAGAGAGAGCTCTGCTTTCTCCGGCCAAATGATGC 1707
Qy	419	AAACAGGTTTCGTCATTATACAAAATTAATCATATTTTAAAGGAAGCAAAATATAA 478
Db	1708	AAGACCGGCTTTCGCAATTGTACAAAGTACCGCGGTTTAAATCCCAAGGCTACGAT 1767
Qy	479	CGATCCAGTGGTGGGCTGCTCTCCAAAGTGATTTCAAGTGCTCTATCAAAGAGGAGATA 538
Db	1768	TGGAGTGAGCCCTTCAGCCCCGGGGAAGATGAATTTAAGTGCCCATTTAAGGAAGAGAT 1827
Qy	539	GCAATTACCAAGTGGTGAATGGGAAGTTCTTGGCCGGCATGGATCTAATATCCAAGTTGAT 598
Db	1828	GCTCTGACAGCGGTGAATGGGAGGTTTTTGGCGAGGCACGGCTCCAAGATCTGGGTCAAT 1887
Qy	599	GAAGTCAGAAGGCTGGTATATTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACCTG 658
Db	1888	GAGGAGACCAAGCTGGTACTTCCAGGGCACCAAGGACACCGCGCTGGAGCACCACTC 1947
Qy	659	TAGTAGTCAGTTACGTAAATCTGGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCA 718
Db	1948	TAGTGGTCAGCTATGAGGGCGCCGCGAGATCGTACGCCCTCACACGCCCGGCTTCTCC 2007
Qy	719	CATTCTTGTGTCATCAGTCAGCACTGTGACTTCTTTATAAGTAAGTATAGTAACCAAG 778
Db	2008	CATAGCTGCTCCATGAGCCAGAACTTCGACATGTTTCGTACGCCACTACAGCAGCGTGA 2067
Qy	779	AATCCACACTGTGTGCTTTCATAGCTATCAAGTCTCTGAAGTACCCAACTTGCATA 838
Db	2068	ACGCGGCTTGGCTGACGCTTACAAGCTGAGCGGCCCGGACGAGACCCCTCTGCACAG 2127
Qy	839	ACAAGGAATTTGGGGCACCATTTGGATTGAGCAGTCTCTTCTCTGACTATCTCT 898
Db	2128	CAGCCCGCTTCTGGGCTAGCATGATGGAGCGAGCGAGCTGCCCGCGGATTTATGTTCT 2187
Qy	899	CCAGAAATTTCTCTTTGAAAGTACTACTGGAATTAATTTGATGGGATGCTCTACAAG 958
Db	2188	CCAGAGATCTTCCATTTCCACACGCGCTCGGATGTGCGGCTCTACGGCATGATCTACAAG 2247

Qy	959	CCTCATGATCTACAGCCTGGAAGAAATATCCTACTGTGCTGTTTCATATATGGTGTCT 1018
Db	2248	CCCCACGCCTTGCAGCCAGGAAGACACCCACCCCTCTTTGTATATGGAGGGCCC 2307
Qy	1019	CAGGT 1023
Db	2308	CAGGT 2312

Search completed: May 3, 2006, 02:12:47
Job time : 547.632 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 4, 2006, 03:41:02 ; Search time 1484.73 Seconds
(without alignments)
9295.694 Million cell updates/sec

Title: US-10-825-632-6
Perfect score: 1669
Sequence: 1 acacagtcacgaatccta.....aaaaaaaaaaaaaaaaaaaaa 1669

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134699005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA Main:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1669	100.0	1669	8	US-10-825-632-6
2	1667	99.9	2830	9	US-10-956-157-2177
3	1648.4	98.8	4523	3	US-09-976-674-8
4	1648.4	98.8	4523	9	US-10-982-512-8
5	1501	89.9	4676	3	US-09-976-674-20
6	1501	89.9	4676	9	US-10-982-512-20
7	1353.4	81.1	3143	6	US-10-170-789-37
8	1349.6	80.9	4829	3	US-09-976-674-12
9	1349.6	80.9	4829	9	US-10-982-512-12
10	1347	80.7	3120	7	US-10-415-122-5
11	1347	80.7	3120	8	US-10-825-632-2
12	1344	80.5	3106	7	US-10-311-035-30
13	1343.6	80.5	4685	3	US-09-976-674-22
14	1343.6	80.5	4685	9	US-10-982-512-22
15	1250.6	74.9	2510	7	US-10-275-505-16
16	1250.6	74.9	2510	10	US-11-140-224-16
17	1087.4	65.2	2649	6	US-10-170-789-39
18	1085.8	65.1	2649	3	US-10-054-776-1
19	1085.8	65.1	2671	3	US-09-976-674-2
20	1085.8	65.1	2671	9	US-10-982-512-2
21	881	52.8	1197	8	US-10-825-632-4
22	765.2	45.8	4309	3	US-09-976-674-14
23	765.2	45.8	4309	9	US-10-982-512-14

24	636	38.1	925	6	US-10-264-237-710	Sequence 710, Appl
25	600	35.9	600	9	US-10-956-157-7412	Sequence 7412, Ap
26	434.2	26.0	1083	8	US-10-825-632-8	Sequence 8, Appli
27	321	19.2	3287	7	US-10-415-122-3	Sequence 3, Appli
28	313	18.8	2495	7	US-10-415-122-8	Sequence 8, Appli
29	313	18.8	2617	3	US-09-976-674-4	Sequence 4, Appli
30	313	18.8	2617	9	US-10-982-512-4	Sequence 223, App
31	313	18.8	2660	7	US-10-072-012-223	Sequence 225, App
32	313	18.8	2660	7	US-10-072-012-225	Sequence 1, Appli
33	313	18.8	3000	7	US-10-415-122-1	Sequence 30, Appl
34	313	18.8	3716	9	US-10-433-757-30	Sequence 32, Appl
35	313	18.8	4076	3	US-09-976-674-32	Sequence 32, Appl
36	313	18.8	4076	9	US-10-982-512-32	Sequence 30, Appl
37	313	18.8	4159	3	US-09-976-674-30	Sequence 30, Appl
38	313	18.8	4159	9	US-10-982-512-30	Sequence 28, Appl
39	313	18.8	4219	3	US-09-976-674-28	Sequence 28, Appl
40	313	18.8	4219	9	US-10-982-512-28	Sequence 24, Appl
41	313	18.8	4302	3	US-09-976-674-24	Sequence 24, Appl
42	313	18.8	4302	9	US-10-982-512-24	Sequence 47, Appl
43	255	15.3	2261	6	US-10-094-749-47	Sequence 40, Appl
44	250.6	15.0	4037	3	US-09-976-674-40	Sequence 40, Appl
45	250.6	15.0	4037	9	US-10-982-512-40	Sequence 40, Appl

ALIGNMENTS

RESULT 1
US-10-825-632-6
; Sequence 6, Application US/10825632
; Publication No. US20040191826A1
; GENERAL INFORMATION:
; APPLICANT: ABBOTT, Catherine Anne
; APPLICANT: GORRELL, Mark Douglas
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES
; FILE REFERENCE: FCSB-100-Div. 1
; CURRENT APPLICATION NUMBER: US/10/825,632
; CURRENT FILING DATE: 2004-04-15
; PRIOR APPLICATION NUMBER: US 10/070,464
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: PCT/AU00/01085
; PRIOR FILING DATE: 2000-09-11
; PRIOR APPLICATION NUMBER: AU PQ5709
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: AU PQ2762
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1669
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-10-825-632-6

Query Match	100.0%;	Score 1669;	DB 8;	Length 1669;
Best Local Similarity	100.0%;	Pred. No. 0;	Mismatches 0;	Indels 0;
Matches 1669;	Conservative 0;			
QY	1	AACAGGTACAGCAAACTCTAAAGTGCATCTTTAAAGATGTCAGAAATAATGATTGCTGA	60	
DB	1	AACAGGTACAGCAAACTCTAAAGTGCATCTTTAAAGATGTCAGAAATAATGATTGCTGA	60	
QY	61	AGGAAGATCATAGATGTCATAGATAAGGAACATAATCAACCTTTGAGATTCATTGGA	120	
DB	61	AGGAAGATCATAGATGTCATAGATAAGGAACATAATCAACCTTTGAGATTCATTGGA	120	
QY	121	AGGAGTTGAATATATTGCCAGAGCTGGATGGACTCCTGAGGGAATAATGCTGGTCCAT	180	
DB	121	AGGAGTTGAATATATTGCCAGAGCTGGATGGACTCCTGAGGGAATAATGCTGGTCCAT	180	
QY	181	CCTACTAGATCGCTCCAGACTCGCCCTACAGATAGTGTGATCTCACCTGAATTTAT	240	
DB	181	CCTACTAGATCGCTCCAGACTCGCCCTACAGATAGTGTGATCTCACCTGAATTTAT	240	

Qy 241 CCAGTAGAAGATGATGTTATATGAAAGGACAGAGACTCATTGAGTCAGTGCCTGATCTGT 300
Db 241 CCAGTAGAAGATGATGTTATGAAAGGACAGAGACTCATTGAGTCAGTGCCTGATCTGT 300
Qy 301 GAGCCCACTAAATATCTATGAGAAACACACAGACATCTGGATAAATATCCATGACATCT 360
Db 301 GAGCCCACTAAATATCTATGAGAAACACACAGACATCTGGATAAATATCCATGACATCT 360
Qy 361 TCATGTTTTTCCCAAGTCAGAAAGGAAATGAGTTATTTTTTAAAGAAAGCAAAATATAACG 420
Db 361 TCATGTTTTTCCCAAGTCAGAAAGGAAATGAGTTATTTTTTAAAGAAAGCAAAATATAACG 420
Qy 421 AACAGTTTCCGTCATTTATACAAAATACATCTATTTTTTAAAGAAAGCAAAATATAACG 480
Db 421 AACAGTTTCCGTCATTTATACAAAATACATCTATTTTTTAAAGAAAGCAAAATATAACG 480
Qy 481 ATCCAGTGTGGGCTGCTGCCATCCAAAGTGATTTCAAAGTGCTCTATCAAAGAGGAGATAGC 540
Db 481 ATCCAGTGTGGGCTGCTGCCATCCAAAGTGATTTCAAAGTGCTCTATCAAAGAGGAGATAGC 540
Qy 541 AATTACAGTGTGTAATGGAAAGTTCTTTGGCCGCGCATGGATCTAATATCCAAGTTGATGA 600
Db 541 AATTACAGTGTGTAATGGAAAGTTCTTTGGCCGCGCATGGATCTAATATCCAAGTTGATGA 600
Qy 601 AGTCAGAAGCTGGTATATTTTGAAGGCACCAAGAGCTCCCTTTAGAGCATCACCTGA 660
Db 601 AGTCAGAAGCTGGTATATTTTGAAGGCACCAAGAGCTCCCTTTAGAGCATCACCTGA 660
Qy 661 CGTAGTCAGTACGTAAATCTCGGAGAGGTGACAAAGCTGACCGTGGCTACTCACA 720
Db 661 CGTAGTCAGTACGTAAATCTCGGAGAGGTGACAAAGCTGACCGTGGCTACTCACA 720
Qy 721 TTCTTCTGCATCAGTCAGCACTGACTTCTTTAAGTAAGTATAGTATACCAAGAA 780
Db 721 TTCTTCTGCATCAGTCAGCACTGACTTCTTTAAGTAAGTATAGTATACCAAGAA 780
Qy 781 TCACACTGTGTGCTCTTACAGCTATCAAGTCTCGAAGATGACCCAACTTGGAAAAC 840
Db 781 TCACACTGTGTGCTCTTACAGCTATCAAGTCTCGAAGATGACCCAACTTGGAAAAC 840
Qy 841 AAGGAATTTTGGGCCACATTTTGGATTACAGAGGTCTCTTCTGACTATACCTCTCC 900
Db 841 AAGGAATTTTGGGCCACATTTTGGATTACAGAGGTCTCTTCTGACTATACCTCTCC 900
Qy 901 AGAAATTTTCTCTTTTGAAGTACTACTGGATTTACATTTGATGGATGCTCTACAAGCC 960
Db 901 AGAAATTTTCTCTTTTGAAGTACTACTGGATTTACATTTGATGGATGCTCTACAAGCC 960
Qy 961 TCATGATCTACAGCCTGGAAAGAAATATCTTACTGTCTGTTTATATATGGTGGTCTCA 1020
Db 961 TCATGATCTACAGCCTGGAAAGAAATATCTTACTGTCTGTTTATATATGGTGGTCTCA 1020
Qy 1021 GGTGTCTATTGTGGGGCCCAAGTCACCTCTGTGGATCTTCTATGATACAGGATACACGA 1080
Db 1021 GGTGTCTATTGTGGGGCCCAAGTCACCTCTGTGGATCTTCTATGATACAGGATACACGA 1080
Qy 1081 AGTTATATGGTGTACCCCTGACCAAGATGACAGGGCTATTTAGGATCTGTGGCCAT 1140
Db 1081 AGTTATATGGTGTACCCCTGACCAAGATGACAGGGCTATTTAGGATCTGTGGCCAT 1140
Qy 1141 SCAGCAGAAAGTTCCCTCTGACCAATTCGTTTACTGCTCTTATCATGGTTTCTCTGA 1200
Db 1141 GCACGAGAAAGTTTCCCTCTGACCAATTCGTTTACTGCTCTTATCATGGTTTCTCTGA 1200
Qy 1201 TGAGATGTCCATTTTGCACATACAGTATATTTACTGAGTTTTTTTGTGAGGGCTGGAAA 1260
Db 1201 TGAGATGTCCATTTTGCACATACAGTATATTTACTGAGTTTTTTTGTGAGGGCTGGAAA 1260
Qy 1261 GCCATATGATTTTACAGATCTATCTCAGGAGAGACACAGCATTAAGAGTTTCTTGAATCGGG 1320
Db 1261 GCCATATGATTTTACAGATCTATCTCAGGAGAGACACAGCATTAAGAGTTTCTTGAATCGGG 1320

Qy 1321 AGACATTATGAACCTGCTATCTTTTGCACCTTCAAGAAAACCTTGGATCACGTATTGC 1380
Db 1321 AGACATTATGAACCTGCTATCTTTTGCACCTTCAAGAAAACCTTGGATCACGTATTGC 1380
Qy 1381 TGCTCTAAAAGTGATATAATTTTACCTGTGTGAGAACTCTCTGGTATACACTGGCTATT 1440
Db 1381 TGCTCTAAAAGTGATATAATTTTACCTGTGTGAGAACTCTCTGGTATACACTGGCTATT 1440
Qy 1441 AACCAATGAGGAGGTTTAAATCAACAGAAAACACAGAAATTGATCATCATCTTTTGATACC 1500
Db 1441 AACCAATGAGGAGGTTTAAATCAACAGAAAACACAGAAATTGATCATCATCTTTTGATACC 1500
Qy 1501 TGCCATGTAACTACTCTCTGAAATATAATGTGTGTCATGCGAGGGTCTACCGTTGT 1560
Db 1501 TGCCATGTAACTACTCTCTGAAATATAATGTGTGTCATGCGAGGGTCTACCGTTGT 1560
Qy 1561 GGTAGTAATCTAATACCTTTAACCCACATGCTCAAAATCAAAATGATACATATCTCTGAGA 1620
Db 1561 GGTAGTAATCTAATACCTTTAACCCACATGCTCAAAATCAAAATGATACATATCTCTGAGA 1620
Qy 1621 GACCCAGCAATACCAATAAGATTTACTAAAAAATAAAAAAAAAAAAAA 1669
Db 1621 GACCCAGCAATACCAATAAGATTTACTAAAAAATAAAAAAAAAAAAAA 1669

RESULT 2
US-10-956-157-2177
; Sequence 2177, Application US/10956157
; Publication No. US20050118625A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William
; TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
; FILE OF INVENTION: HUMAN OSTEOARTHRITIS AND HUMAN PROTEASES
; FILE REFERENCE: 031896-043000 (AM 101081)
; CURRENT APPLICATION NUMBER: US/10/956,157
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 319805
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2177
; LENGTH: 2830
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-956-157-2177

Query Match 99.9%; Score 1667; DB 9; Length 2830;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1667; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AACAGGTACAGCAAAATCTTAAAGTCACCTTTTAAAGATGTCAGAAAATAATGATGCTGA 60
Db 1164 AACAGGTACAGCAAAATCTTAAAGTCACCTTTTAAAGATGTCAGAAAATAATGATGCTGA 1223
Qy 61 AGGAAGGATCATAGATGTCATAGATAAGGAAGTAACTAACTTCAACCTTTTGAGATTCTATTGA 120
Db 1224 AGGAAGGATCATAGATGTCATAGATAAGGAAGTAACTAACTTCAACCTTTTGAGATTCTATTGA 1283
Qy 121 AGGAGTTGAATATATGTCAGAGCTGGATGACCTCTGAGGAAAATATGCTTGGTCCAT 180
Db 1284 AGGAGTTGAATATATGTCAGAGCTGGATGACCTCTGAGGAAAATATGCTTGGTCCAT 1343
Qy 181 CCTACTAGATCCCTCCAGACATCGCCCTACAGATAGTGTGATCTCACCTGAAATATTAT 240
Db 1344 CCTACTAGATCCCTCCAGACATCGCCCTACAGATAGTGTGATCTCACCTGAAATATTAT 1403
Qy 241 CCAGTGAAGATGATGTTATGGAAGGACAGAGACTCAATTGAGTCAGTGCCTGATCTGT 300
Db 1404 CCAGTGAAGATGATGTTATGGAAGGACAGAGACTCAATTGAGTCAGTGCCTGATCTGT 1463
Qy 301 GAGCCCACTAAATTTATCTATGAAGAAACACACAGACATCTGGATAAATATCCATGACATCT 360
Db 1464 GAGCCCACTAAATTTATCTATGAAGAAACACACAGACATCTGGATAAATATCCATGACATCT 1523

QY 361 TCATGTTTTTCCCAAGTACGAGAGAAATTTGAGTTTATTTTTTTCCTCTCGAATGCAA 420
Db 1524 TCATGTTTTTCCCAAGTACGAGAGAAATTTGAGTTTATTTTTTTCCTCTCGAATGCAA 1583
QY 421 AACAGGTTTCCGTCATTTATACAAATTTACATCTATTTTAAAGGAAAGCAAAATATAACG 480
Db 1584 AACAGGTTTCCGTCATTTATACAAATTTACATCTATTTTAAAGGAAAGCAAAATATAACG 1643
QY 481 ATCCAGTGTGGGTCCTGCTCCCAAGTGATTTCAAGTGTCTTCAAGTGTCTTCAAGAGGAGATAGC 540
Db 1644 ATCCAGTGTGGGTCCTGCTCCCAAGTGATTTCAAGTGTCTTCAAGAGGAGATAGC 1703
QY 541 AATACCACTGCTGATGGAGTTCTTTGGCGGCATGGATCTAATATCAAGTTGATGA 600
Db 1704 AATACCACTGCTGATGGAGTTCTTTGGCGGCATGGATCTAATATCAAGTTGATGA 1763
QY 601 AGTCAGAGGCTGGTATATTTTGAAGGCACCAAGACTCTCCCTTTAGAGCATCACCTGTA 660
Db 1764 AGTCAGAGGCTGGTATATTTTGAAGGCACCAAGACTCTCCCTTTAGAGCATCACCTGTA 1823
QY 661 CGTAGTCAGTTACGTAATTCCTGGAGAGGTGACAAAGGCTGACGACCGTGGCTACTACA 720
Db 1824 CGTAGTCAGTTACGTAATTCCTGGAGAGGTGACAAAGGCTGACGACCGTGGCTACTACA 1883
QY 721 TTCTTGCTGCATCAGTCAGACTGACCTCTTTATAGTAAGTATAGTAACCAAGAGAA 780
Db 1884 TTCTTGCTGCATCAGTCAGACTGACCTCTTTATAGTAAGTATAGTAACCAAGAGAA 1943
QY 781 TCCACACTGTGTCTCTTTCAAGCTATCAAGTCTTGAAGATGACCCAACTTGCAGAAC 840
Db 1944 TCCACACTGTGTCTCTTTCAAGCTATCAAGTCTTGAAGATGACCCAACTTGCAGAAC 2003
QY 841 AAAGGAATTTGGGCGACCATTTGGATTGACGAGGTCTCTCTCTGACTATACTCTCC 900
Db 2004 AAAGGAATTTGGGCGACCATTTGGATTGACGAGGTCTCTCTCTGACTATACTCTCC 2063
QY 901 AGAAATTTCTCTTTGAAAGTACTACTGGAATTTACATTTGTATGGGATGCTCTACAAGC 960
Db 2064 AGAAATTTCTCTTTGAAAGTACTACTGGAATTTACATTTGTATGGGATGCTCTACAAGC 2123
QY 961 TCATGATCTACAGCTGGAAGAAATATCTCTGCTGTCTGTCTCATATATAGTGGTCTCA 1020
Db 2124 TCATGATCTACAGCTGGAAGAAATATCTCTGCTGTCTGTCTCATATATAGTGGTCTCA 2183
QY 1021 GGTGCTATTTGCTGGGCGCCAGTCACTCTGTGATCTTCTATGATACAGGATACACGA 1080
Db 2184 GGTGCTATTTGCTGGGCGCCAGTCACTCTGTGATCTTCTATGATACAGGATACACGA 2243
QY 1081 ACCTTATATGGGTCACTCCCTGACCAAGATGAACAGGCTATTTACTTAGGATCTGTGGCCAT 1140
Db 2244 ACCTTATATGGGTCACTCCCTGACCAAGATGAACAGGCTATTTACTTAGGATCTGTGGCCAT 2303
QY 1141 GCAGCAGAAAAGTTCCCTCTGAAACAAATCGTTTACTGCTCTTACATGTTTCTCTGGA 1200
Db 2304 GCAGCAGAAAAGTTCCCTCTGAAACAAATCGTTTACTGCTCTTACATGTTTCTCTGGA 2363
QY 1201 TGAGATGTCATTTTGCATACACAGTATATTTACTGATTTTCTAGTAGGGCTGGA 1260
Db 2364 TGAGATGTCATTTTGCATACACAGTATATTTACTGATTTTCTAGTAGGGCTGGA 2423
QY 1261 GCCATATGATTTACAGATCTATCTCTCAGGAGAGACACAGATAGAGTTCTCTGAATCGGG 1320
Db 2424 GCCATATGATTTACAGATCTATCTCTCAGGAGAGACACAGATAGAGTTCTCTGAATCGGG 2483
QY 1321 AGAACATTAATGATGCTATCTTTGCACTACCTTTCAAGAAAACCTTTGGATCAGTATTCG 1380
Db 2484 AGAACATTAATGATGCTATCTTTGCACTACCTTTCAAGAAAACCTTTGGATCAGTATTCG 2543
QY 1381 TGTCTTAAAGTGATATAATTTGACCTGTGTAGAACTCTCTGCTATACACTGGCTATTT 1440
Db 2544 TGTCTTAAAGTGATATAATTTGACCTGTGTAGAACTCTCTGCTATACACTGGCTATTT 2603
QY 1441 AACCAATGAGGAGGTTTTTAATCAACAGAAAAACAGAAATTTGATCATCATCTTTTGATACC 1500

Db 2604 AACCAATGAGGAGGTTTTTAATCAACAGAAAAACAGAAATTTGATCATCATTTTGATACC 2663
QY 1501 TGCATGTAACTACTACTCTCTGAAATAAATGCTGCTGCATGCGAGGGGTCTACCGTTTGT 1560
Db 2664 TGCATGTAACTACTACTCTCTGAAATAAATGCTGCTGCATGCGAGGGGTCTACCGTTTGT 2723
QY 1561 GGTAGTAATCTAATACCTTTAAACCCACATGCTCAAAATCAAAATGATACATATCTCTGAGA 1620
Db 2724 GGTAGTAATCTAATACCTTTAAACCCACATGCTCAAAATCAAAATGATACATATCTCTGAGA 2783
QY 1621 GACCCAGCAATACCAATGAATTTACTTAAATAAAAAAAAAAAAAA 1667
Db 2784 GACCCAGCAATACCAATGAATTTACTTAAATAAAAAAAAAAAAAA 2830
RESULT 3
US-09-976-674-8
; Sequence 8, Application US/09976674
; Patent No. US20020115943A1
; GENERAL INFORMATION:
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 4523
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-976-674-8
Query Match 98.8%; Score 1648.4; DB 3; Length 4523;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1663; Conservative 0; Mismatches 1; Indels 4; Gaps 1;
QY 1 AACAGGTACAGCAAAATCTTAAAGTCACTTTTAAAGATGTCAGAAATAATGATGATGCTGA 60
Db 1164 AACAGGTACAGCAAAATCTTAAAGTCACTTTTAAAGATGTCAGAAATAATGATGATGCTGA 1223
QY 61 AGGAAGGATCATAGATGTCATAGATAAGAACTAAATCAACCTTTTGAGATTCCTATTGA 120
Db 1224 AGGAAGGATCATAGATGTCATAGATAAGAACTAAATCAACCTTTTGAGATTCCTATTGA 1283
QY 121 AGGAGTTGAATATATGCGCAGAGCTGGATGGACTCTGAGGGAATAATGCTGGTCCAT 180
Db 1284 AGGAGTTGAATATATGCGCAGAGCTGGATGGACTCTGAGGGAATAATGCTGGTCCAT 1343
QY 181 CCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAATTTATTAT 240
Db 1344 CCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAATTTATTAT 1403
QY 241 CCAGTAGAAGATGATGTTATGGAAGGAGAGACTCAATTGAGTCAGTCGCTGATTCGT 300
Db 1404 CCAGTAGAAGATGATGTTATGGAAGGAGAGACTCAATTGAGTCAGTCGCTGATTCGT 1463
QY 301 GAGCCACATAATTTCTATGAAGAAACACAGACATCTGGATAAATATCCATGACATCTT 360
Db 1464 GAGCCACATAATTTCTATGAAGAAACACAGACATCTGGATAAATATCCATGACATCTT 1523
QY 361 TCATGTTTTTCCCAAGTACAGAGAGAAATTTGAGTTTATTTTTTGCCTCTGAATGCAA 420
Db 1524 TCATGTTTTTCCCAAGTACAGAGAGAAATTTGAGTTTATTTTTTGCCTCTGAATGCAA 1583
QY 421 AACAGGTTTCCGTCATTTATACAAATTTACATCTATTTTAAAGGAAAGCAAAATATAACG 480

Db 1584 AACAGTTTCCGTCATTATACAAATTAACATCTATTTTAAAGGAAACCAATATATAACG 1643
Qy 481 ATCAAGTGTGGGCTGCCTCCCAAGTGAATTTCAAGTGTCTTCAAGAGGAGATAGC 540
Db 1644 ATCCAGTGTGGGCTGCCTCCCAAGTGAATTTCAAGTGTCTTCAAGAGGAGATAGC 1703
Qy 541 AATTACAGTGTGTAATGGAAATTTCTTGGCGGCGATGGATCTAATATCCAAGTTCATGA 600
Db 1704 AATTACAGTGTGTAATGGAAATTTCTTGGCGGCGATGGATCTAATATCCAAGTTCATGA 1763
Qy 601 AGTCAGAAGGCTGGTATATTTTGAAGGCACCAAGAGCTCCCTTTTAGAGCATCACTGTA 560
Db 1764 AGTCAGAAGGCTGGTATATTTTGAAGGCACCAAGAGCTCCCTTTTAGAGCATCACTGTA 1823
Qy 661 CGTAGTCAGTTACGTAATCTTGGAGAGGTGACAAAGCTGACACGGTGGCTACTACA 720
Db 1824 CGTAGTCAGTTACGTAATCTTGGAGAGGTGACAAAGCTGACACGGTGGCTACTACA 1883
Qy 721 TTCTTGCTGCATCAGTCAGCACTGTGACTTCTTTAAGTAAGTATAGTAACCAAGAA 780
Db 1884 TTCTTGCTGCATCAGTCAGCACTGTGACTTCTTTAAGTAAGTATAGTAACCAAGAA 1943
Qy 781 TCCACACTGTGTGCTTCCCTTACAAAGCTATCAAGTCTGGAAGATGACCCAACTTGCAAAAC 840
Db 1944 TCCACACTGTGTGCTTCCCTTACAAAGCTATCAAGTCTGGAAGATGACCCAACTTGCAAAAC 2003
Qy 841 RAAGGAATTTGGGCCACCAATTTTGGATTCAGAGGTCCTTCTTCGACTACTCTCC 900
Db 2004 RAAGGAATTTGGGCCACCAATTTTGGATTCAGAGGTCCTTCTTCGACTACTCTCC 2063
Qy 901 AGAAATTTTCTTTTGAAGTACTACTGGATTACATTTGATGGGATGCTCTACAAGCC 960
Db 2064 AGAAATTTTCTTTTGAAGTACTACTGGATTACATTTGATGGGATGCTCTACAAGCC 2123
Qy 961 TCATGATCTACAGCTCGAAAGAAATATCTTACTGTGCTGTTTCATATATGGTGGTCTCA 1020
Db 2124 TCATGATCTACAGCTCGAAAGAAATATCTTACTGTGCTGTTTCATATATGGTGGTCTCA 2179
Qy 1021 GGTGTCTATTTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTTCTGACTACTCTCC 1080
Db 2180 GGTGTCTATTTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTTCTGACTACTCTCC 2239
Qy 1081 ACGTATATGGGTCAACCTTACACAGAAATGAAGGCTATTAATCTAGGATCTGGGCCAT 1140
Db 2240 ACGTATATGGGTCAACCTTACACAGAAATGAAGGCTATTAATCTAGGATCTGGGCCAT 2299
Qy 1141 GCAAGCAGAAAGTTCCCTCTGAAACCAATCGTTTACTGCTTCTTACATGGTTCCTGGA 1200
Db 2300 GCAAGCAGAAAGTTCCCTCTGAAACCAATCGTTTACTGCTTCTTACATGGTTCCTGGA 2359
Qy 1201 TGAGATGTCCATTTTGCACATACAGTATATTAAGTGTGTTTCTAGTGGGCTGGAA 1260
Db 2360 TGAGATGTCCATTTTGCACATACAGTATATTAAGTGTGTTTCTAGTGGGCTGGAA 2419
Qy 1261 GCCATATGATTTACAGATCTATCTCAGGAGACACAGCATAGAGTTCTGATCGGG 1320
Db 2420 GCCATATGATTTACAGATCTATCTCAGGAGACACAGCATAGAGTTCTGATCGGG 2479
Qy 1321 AGACATATGAATCGATCTTTTGGACCTTCTGACATCTTCAAGAAACCTTGGATCACTATTC 1380
Db 2480 AGACATATGAATCGATCTTTTGGACCTTCTGACATCTTCAAGAAACCTTGGATCACTATTC 2539
Qy 1381 TGCTCTAAAGTGAATATTTTGAACCTGTGTAGAACTCTCTGGTATACACTGGCTATTT 1440
Db 2540 TGCTCTAAAGTGAATATTTTGAACCTGTGTAGAACTCTCTGGTATACACTGGCTATTT 2599
Qy 1441 AACCAATAGAGGTTTAAATCAACAGAAAAACAGAAATGATCATCACATTTTGTATACC 1500
Db 2600 AACCAATAGAGGTTTAAATCAACAGAAAAACAGAAATGATCATCACATTTTGTATACC 2659
Qy 1501 TGCCATGTAACTACTCTCTGAAATTAATATGGTGGCCATGCGAGGGGTCTACGGTTTGT 1560
Db 2660 TGCCATGTAACTACTCTCTGAAATTAATATGGTGGCCATGCGAGGGGTCTACGGTTTGT 2719

Qy 1561 GGTAGTAATCTATACCTTAACCCACATGCTCAAAATCAATCATATTCCTCAGA 1620
Db 2720 GGTAGTAATCTATACCTTAACCCACATGCTCAAAATCAATCATATTCCTCAGA 2779
Qy 1621 GACCCAGCAATACCATAGAAATTAATAAAAAAAAAAAAAAAAAAAAAA 1668
Db 2780 GACCCAGCAATACCATAGAAATTAATAAAAAAAAAAAAAAAAAAAAAA 2827

RESULT 4
US-10-982-512-8
; Sequence 8, Application US/10982512
; Publication No. US20050059081A1
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/10/982,512
; PRIOR FILING DATE: 2004-11-05
; PRIOR APPLICATION NUMBER: US/09/976,674
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 4523
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-982-512-8

Query Match 98.8%; Score 1648.4; DB 9; Length 4523;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1663; Conservative 0; Mismatches 1; Indels 4; Gaps 1;

Qy 1 AACAGTACAGCAATCTTAAAGTCACTTTTAAGATGTGCAAAATTAATGATGCTGA 60
Db 1164 AACAGTACAGCAATCTTAAAGTCACTTTTAAGATGTGCAAAATTAATGATGCTGA 1223
Qy 61 AGGAAGGATCATAGATGTCTAGATAAGGAACATAATCAACCTTTTGAGATCTATTGA 120
Db 1224 AGGAAGGATCATAGATGTCTAGATAAGGAACATAATCAACCTTTTGAGATCTATTGA 1283
Qy 121 AGGAGTTGAATATATTCACAGAGCTGGATGGACTCCTGAGGGAATAATGCTTGGTCCAT 180
Db 1284 AGGAGTTGAATATATTCACAGAGCTGGATGGACTCCTGAGGGAATAATGCTTGGTCCAT 1343
Qy 181 CCTACTAGATCGCTCCAGAGCTCGCCTTACAGATAGTGTGATCTCACCTGAATATTAT 240
Db 1344 CCTACTAGATCGCTCCAGAGCTCGCCTTACAGATAGTGTGATCTCACCTGAATATTAT 1403
Qy 241 CCCAGTAGAAGATGATGTTATGGAAGGCGAGAGACTCATTGAGTCAGTGCCCTGATCTGT 300
Db 1404 CCCAGTAGAAGATGATGTTATGGAAGGCGAGAGACTCATTGAGTCAGTGCCCTGATCTGT 1463
Qy 301 GACGCCACTAATATCTATGAAGAAACAAACAGACATCTGGAATAATATCCATGACATCTT 360
Db 1464 GACGCCACTAATATCTATGAAGAAACAAACAGACATCTGGAATAATATCCATGACATCTT 1523
Qy 361 TCATGTTTTTCCCCAAAGTCACGAAGAGAAATTTGAGTTTATTTTTGGCTCTGATGCAA 420
Db 1524 TCATGTTTTTCCCCAAAGTCACGAAGAGAAATTTGAGTTTATTTTTGGCTCTGATGCAA 1583
Qy 421 AACAGGTTTCCGTCATTATACAAAATTAATCTATTTTAAGGAAGCAATATAACG 480
Db 1584 AACAGGTTTCCGTCATTATACAAAATTAATCTATTTTAAGGAAGCAATATAACG 1643
Qy 481 ATCCAGTGGTGGGCTCCCTGCTCCAGTGAATTTCAAGTGTCTCTATCAAGAGGATAGC 540

1644 ATCCAGTGGTGGGCTGCTGCCAAGTGAATTTCAAGTGTCTTCAAGAGGAGATAGC 1703
541 AATTACAGTGGTGAATGGGAAGTTCTTGGCGGCATGGATCTAATATCAAGTTGATGA 600
1704 AATTACAGTGGTGAATGGGAAGTTCTTGGCGGCATGGATCTAATATCAAGTTGATGA 1763
601 AGTCAGAGGCTGGTATATTTTGAAGGCACAAAGACTCCCTTTAGAGCATCACCTGTGA 660
1764 AGTCAGAGGCTGGTATATTTTGAAGGCACAAAGACTCCCTTTAGAGCATCACCTGTGA 1823
661 CGTAGTCAGTTACGTAAATCCCTGGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACA 720
1824 CGTAGTCAGTTACGTAAATCCCTGGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACA 1883
721 TTCTTCTGTCATCAGTCAGCATCTGTGACTTCTTTATAGTAAGTATAGTAAACAGAGAA 780
1884 TTCTTCTGTCATCAGTCAGCATCTGTGACTTCTTTATAGTAAGTATAGTAAACAGAGAA 1943
781 TCACACTGTGTGCTCTTTTCAAGCTATCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCT 840
1944 TCACACTGTGTGCTCTTTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCT 2003
841 AAGGAATTTTGGGCCACCAATTTGGATTTCAGAGTTCAGAGTTCAGAGTTCAGAGTTCAG 900
2004 AAGGAATTTTGGGCCACCAATTTGGATTTCAGAGTTCAGAGTTCAGAGTTCAGAGTTCAG 2063
901 AGAAATTTCTCTTTTGAAGTACTTGGATTGTATGATTTGATGGGATGCTTACAAAGCC 960
2064 AGAAATTTCTCTTTTGAAGTACTTGGATTGTATGATTTGATGGGATGCTTACAAAGCC 2123
961 TCATGATCTACAGCTTGGAAAGAAATATCTTACTGTGCTGTGTTTATATATGTTGCTCA 1020
2124 TCATGATCTACAGCTTGGAAAGAAATATCTTACTGTGCTGTGTTTATATATGTTGCTC 2179
1021 GGTGTGCTATGTGCGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACAGGA 1080
2180 GGTGTGCTATGTGCGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACAGGA 2239
1081 AGTTATATGGGTCACTTACAGATGACAGATGACAGGCTATTTACTTAGGATCTGTGGCCAT 1140
2240 AGTTATATGGGTCACTTACAGATGACAGATGACAGGCTATTTACTTAGGATCTGTGGCCAT 2299
1141 GCAAGCAGAAAGTTTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGATGTTTCTCGGA 1200
2300 GCAAGCAGAAAGTTTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGATGTTTCTCGGA 2359
1201 TGAGATGTCATTTTGCACATACAGATATATTTACTGAGTGTGTTTGTAGGGGCTGGAAA 1260
2360 TGAGATGTCATTTTGCACATACAGATATATTTACTGAGTGTGTTTGTAGGGGCTGGAAA 2419
1261 GCCATATGATTTACAGATCTATCTCAGGAGAGACACAGATAGAGTTTCTGAATCGGG 1320
2420 GCCATATGATTTACAGATCTATCTCAGGAGAGACACAGATAGAGTTTCTGAATCGGG 2479
1321 AGAACATATGAACTGTCATCTTTTGCATPACCTTCAAGAAAACCTTGGATCAGATATGC 1380
2480 AGAACATATGAACTGTCATCTTTTGCATPACCTTCAAGAAAACCTTGGATCAGATATGC 2539
1381 TGCTCTAAAGTGAATATATTTTGAACCTGTGTGAGAACTCTCTGGTATACACTGGCTATTT 1440
2540 TGCTCTAAAGTGAATATATTTTGAACCTGTGTGAGAACTCTCTGGTATACACTGGCTATTT 2599
1441 AACCAATGAGGAGTTTAAATCAACAGAAAACACAGATTTGATCATCATTTTGTATACC 1500
2600 AACCAATGAGGAGTTTAAATCAACAGAAAACACAGATTTGATCATCATTTTGTATACC 2659
1501 TGCCATGTACATCTACTCTGAAATAAATGCTGTGTCATGCTGAGGGGCTTACGGTTTGT 1560
2660 TGCCATGTACATCTACTCTGAAATAAATGCTGTGTCATGCTGAGGGGCTTACGGTTTGT 2719
1561 GGTAGTAACTTAATCTTAAACCCCATGCTCAAAATCAAAATGATATATCTTCTGAGA 1620
2720 GGTAGTAACTTAATCTTAAACCCCATGCTCAAAATCAAAATGATATATCTTCTGAGA 2779

QY 1621 GACCCAGCATACCATAGAAATTTACTTAAATAAAAAAAAAAAAAAAAAAAAA 1668
Db 2780 GACCCAGCATACCATAGAAATTTACTTAAATAAAAAAAAAAAAAAAAAAAAA 2827
RESULT 5
US-09-976-674-20
; Sequence 20, Application US/09976674
; Patent No. US20020115843A1
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20
; LENGTH: 4676
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-976-674-20
Query Match 89.9%; Score 1501; DB 3; Length 4676;
Best Local Similarity 91.5%; Pred. No. 0;
Matches 1663; Conservative 0; Mismatches 5; Indels 149; Gaps 1;
QY 1 AACAGGTACAGCAAAATCCCTAAAGTCACTTTTAAGATGTGAGAAATAAATGATGCTGA 60
Db 1164 AACAGGTACAGCAAAATCCCTAAAGTCACTTTTAAGATGTGAGAAATAAATGATGCTGA 1223
QY 61 AGGAAGATCATAGATGTATAGATAAGAACTAAATTCACCTTTTGAAGATTTCTATTGA 120
Db 1224 AGGAAGATCATAGATGTATAGATAAGAACTAAATTCACCTTTTGAAGATTTCTATTGA 1283
QY 121 AGGAGTTGAATATATTGCCAGAGCTGCGCTACAGATAGTGTGATGCTGCTGATTCCT 180
Db 1284 AGGAGTTGAATATATTGCCAGAGCTGCGCTACAGATAGTGTGATGCTGCTGATTCCT 1343
QY 181 CCTACTAGATGCTGCCAGACTGCGCTACAGATAGTGTGATGCTGCTGATTCCTGATTCCT 240
Db 1344 CCTACTAGATGCTGCCAGACTGCGCTACAGATAGTGTGATGCTGCTGATTCCTGATTCCT 1403
QY 241 CCCAGTGAAGATGATGTTATGGAAGGAGAGACTCATTGAGTCAGTGCCTGATTCCTGT 300
Db 1404 CCCAGTGAAGATGATGTTATGGAAGGAGAGACTCATTGAGTCAGTGCCTGATTCCTGT 1463
QY 301 GACCCCACTAATTTCTATGAAGAAACACAGACATCTGGAATAATATCCATGACATCTT 360
Db 1464 GACCCCACTAATTTCTATGAAGAAACACAGACATCTGGAATAATATCCATGACATCTT 1523
QY 361 TCATGTTTTTCCCCAAAGTCACGAAGAGAAATGAGTTTATTTTGGCTCTGAAATGCAA 420
Db 1524 TCATGTTTTTCCCCAAAGTCACGAAGAGAAATGAGTTTATTTTGGCTCTGAAATGCAA 1583
QY 421 AACAGGTTTCGGTCAATTTATACAAAATTTACATCTATTTTAAAGGAAAGCAATATAAAC 480
Db 1584 AACAGGTTTCGGTCAATTTATACAAAATTTACATCTATTTTAAAGGAAAGCAATATAAAC 1643
QY 481 ATCCAGTGGTGGGCTGCTGCTCCAGTGAATTTCAAGTGTCTTCAAGAGGAGATAGC 540
Db 1644 ATCCAGTGGTGGGCTGCTGCTCCAGTGAATTTCAAGTGTCTTCAAGAGGAGATAGC 1703
QY 541 AATTACAGTGGTGAATGGGAAGTTCTTGGCCGGCATGGAATCTAATATCCAGTTGATGA 600
Db 1704 AATTACAGTGGTGAATGGGAAGTTCTTGGCCGGCATGGAATCTAATATCCAGTTGATGA 1763

QY 601 AGTCAGAGGCTGGTATATTTTGAAGGCACCAAGGACTCCCTTTAGAGCATCACCTGTA 660
Db 1764 AGTCAGAGGCTGGTATATTTTGAAGGCACCAAGGACTCCCTTTAGAGCATCACCTGTA 1823
QY 661 GGTAGTCAGTTACGTAATCTCTGGAGAGGTGACAGGCTGACTGACCGTGGCTACTCACA 720
Db 1824 CGTAGTCAGTTACGTAATCTCTGGAGAGGTGACAGGCTGACTGACCGTGGCTACTCACA 1883
QY 721 TTCTTGCTGATCAGTCAGACTGTGACTTCTTTAAGTAAGTATAGTAACGAGAGAA 780
Db 1884 TTCTTGCTGATCAGTCAGACTGTGACTTCTTTAAGTAAGTATAGTAACGAGAGAA 1943
QY 781 TCACACTGTGTGCTCCCTTTACAAGCTATCAAGTCTGAAGTCAACCACTTGCAAAAC 840
Db 1944 TCACACTGTGTGCTCCCTTTACAAGCTATCAAGTCTGAAGTCAACCACTTGCAAAAC 2003
QY 841 AAGGAAATTTTGGGCCACCAATTTTGGATTTCAGCAGGTCCTCTTCTGACTATATCTCTCC 900
Db 2004 AAGGAAATTTTGGGCCACCAATTTTGGATTTCAGCAGGTCCTCTTCTGACTATATCTCTCC 2063
QY 901 AGAAATTTTCTCTTTTGAAGGACTACTGGAATTTACATTTGATGGGATGCTCTACAAGCC 960
Db 2064 AGAAATTTTCTCTTTTGAAGGACTACTGGAATTTACATTTGATGGGATGCTCTACAAGCC 2123
QY 961 TCATGATCTACAGCCTGGAAAGAAATATCCTACTGTGCTGTTCATATATGTTGTC--- 1016
Db 2124 TCATGATCTACAGCCTGGAAAGAAATATCCTACTGTGCTGTTCATATATGTTGTCGGGT 2183
QY 1017 ----- 1016
Db 2184 CAAATAGAAATGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGTTTC 2243
QY 1017 ----- 1016
Db 2244 ATTGACTTAGATCGTGGGCATCCAGGCTGGTCCCTATGGAGGATACCTTCCCTGATG 2303
QY 1017 -----CTCAGGTTCTGATTTCTGGGCCCCAGTCACTCTG 1051
Db 2304 GCATTAAATCGAGGTCAGATATCTTACGGGTGCTATGCTGGGCCCCAGTCACTCTG 2363
QY 1052 TGGATCTCTATGATACAGGATACAGGAAAGTTATATGGTTCACCTGACCAAGATGAA 1111
Db 2364 TGGATCTCTATGATACAGGATACAGGAAAGTTATATGGTTCACCTGACCAAGATGAA 2423
QY 1112 CAGGCTATTTACTTAGGATCTGTGGCCATGCAAGCAGAAAGTTTCCCTCTGAAACCAAT 1171
Db 2424 CAGGCTATTTACTTAGGATCTGTGGCCATGCAAGCAGAAAGTTTCCCTCTGAAACCAAT 2483
QY 1172 CGTTTACTGCTCTTACATGGTTTCTGGATGAGAAATGTCATTTTGACATACAGATATA 1231
Db 2484 CGTTTACTGCTCTTACATGGTTTCTGGATGAGAAATGTCATTTTGACATACAGATATA 2543
QY 1232 TTACTGAGTTTTTATGAGGCTGGAAGCCATATGATTTACAGATCTATCTCTCAGGAG 1291
Db 2544 TTACTGAGTTTTTATGAGGCTGGAAGCCATATGATTTACAGATCTATCTCTCAGGAG 2603
QY 1292 AGACAGCATAGAGTTCTCTGAATCGGGAGAACATTTATGAATGTCATCTTTTGCACTAC 1351
Db 2604 AGACAGCATAGAGTTCTCTGAATCGGGAGAACATTTATGAATGTCATCTTTTGCACTAC 2663
QY 1352 CTTCAAGAAAACCTTTGGATCAGTATTCCTCTCTAAAAGTGATATAATTTTGACCTGTG 1411
Db 2664 CTTCAAGAAAACCTTTGGATCAGTATTCCTCTCTAAAAGTGATATAATTTTGACCTGTG 2723
QY 1412 TAGAACTCTCTGTTATACACTGGCTATTTTAACCAATGAGAGGTTTAAATCAACAGAAA 1471
Db 2724 TAGAACTCTCTGTTATACACTGGCTATTTTAACCAATGAGAGGTTTAAATCAACAGAAA 2783
QY 1472 CACAGAAATGATCACAATTTGATACCTGCCATGTAACATCTACTCTGGAATAAAT 1531
Db 2784 CACAGAAATGATCACAATTTGATACCTGCCATGTAACATCTACTCTGGAATAAAT 2843
QY 1532 GTGGTGCCATGCAGGGGCTACGGTTTGTGGTAGTATCTAATACCTTAAACCCCATGC 1591

Db 2844 GTGGTGCCATGCAGGGTCTACGGTTTGTGGTAGTAACTAATACCTTAAACCCCATGC 2903
QY 1592 TCAAAATCAAATGATACATATTCCTGAGAGACCCAGCAATACCATPAAGAATTACTTAAAA 1651
Db 2904 TCAAAATCAAATGATACATATTCCTGAGAGACCCAGCAATACCATPAAGAATTACTTAAAA 2963
QY 1652 AAAAAAAAAAAAAAAAAA 1668
Db 2964 AAAAAAAAAAAAAAAAAA 2980
RESULT 6
US-10-982-512-20
; Sequence 20, Application US/10982512
; Publication No. US20050059081A1
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/10/982,512
; PRIOR FILING DATE: 2004-11-05
; PRIOR APPLICATION NUMBER: US/09/976,674
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20
; LENGTH: 4676
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-982-512-20
Query Match 89.9%; Score 1501; DB 9; Length 4676;
Best Local Similarity 91.5%; Pred. No. 0;
Matches 1663; Conservative 0; Mismatches 5; Indels 149; Gaps 1;
QY 1 AACAGGTACAGCAAAATCCCTAAAGTCACTTTTAAAGATGTCAGAAATATATGATGCTGA 60
Db 1164 AACAGGTACAGCAAAATCCCTAAAGTCACTTTTAAAGATGTCAGAAATATATGATGCTGA 1223
QY 61 AGGAAGGATCATAGATGTTCATAGATAAGGAACATAATCAACCTTTTGAGATTCCTATTGA 120
Db 1224 AGGAAGGATCATAGATGTTCATAGATAAGGAACATAATCAACCTTTTGAGATTCCTATTGA 1283
QY 121 AGGAGTTGAATATATTTGCCAGAGCTGGATGGACTCTCCTGAGGAAAAATATGCTTGGTCCAT 180
Db 1284 AGGAGTTGAATATATTTGCCAGAGCTGGATGGACTCTCCTGAGGAAAAATATGCTTGGTCCAT 1343
QY 181 CCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCAGCTGAATTTATTAT 240
Db 1344 CCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCAGCTGAATTTATTAT 1403
QY 241 CCCAGTAGAAGATGATGTTTATGAAAGGCAGAGACTCAATCAGTCAGTCGCTGATTCGT 300
Db 1404 CCCAGTAGAAGATGATGTTTATGAAAGGCAGAGACTCAATCAGTCAGTCGCTGATTCGT 1463
QY 301 GAGCCCACTAAATTTCTATGAAGAAACCAACAGACATCTGGATAAATATCCATGACATCTT 360
Db 1464 GAGCCCACTAAATTTCTATGAAGAAACCAACAGACATCTGGATAAATATCCATGACATCTT 1523
QY 361 TCATGTTTTTCCCAAGTCCACGAGAGGAATAGTATTATTTTCCCTCTGAATGCAA 420
Db 1524 TCATGTTTTTCCCAAGTCCACGAGAGGAATAGTATTATTTTCCCTCTGAATGCAA 1583
QY 421 AACAGGTTTTCCCTGATTTTATACAAATTTACATCTATTATTAAGGAGAAAGCAATATAACG 480
Db 1584 AACAGGTTTTCCCTGATTTTATACAAATTTACATCTATTATTAAGGAGAAAGCAATATAACG 1643

QY 481 ATCCAGTGGTGGGCTGCTCCTCCAAAGTGATTTCAAGTGTCTATCAAGAGGAGATAGC 540
Db 1644 ATCCAGTGGTGGGCTGCTCCTCCAAAGTGATTTCAAGTGTCTATCAAGAGGAGATAGC 1703
QY 541 AATTACAGTGGTGAATGGGAAGTCTTGGCCGGCATGGATCTAATATCCCAAGTTGATGA 600
Db 1704 AATTACAGTGGTGAATGGGAAGTCTTGGCCGGCATGGATCTAATATCCCAAGTTGATGA 1763
QY 601 AGTCAGAAGGCTGGTATATTTTGAAGGCCACCAAGACTCCCTTTAGAGCATCACTGTGA 660
Db 1764 AGTCAGAAGGCTGGTATATTTTGAAGGCCACCAAGACTCCCTTTAGAGCATCACTGTGA 1823
QY 661 GGTAGTCAGTACGTAATCCTGGAGAGGTGACAGGCTGACTGACCGTGGCTACTCACA 720
Db 1824 GGTAGTCAGTACGTAATCCTGGAGAGGTGACAGGCTGACTGACCGTGGCTACTCACA 1883
QY 721 TTCTTGTGTCATCAGTCAGCACTGTGACTTCTTTTAAAGTAAGTATAGTACCAAGAGAA 780
Db 1884 TTCTTGTGTCATCAGTCAGCACTGTGACTTCTTTTAAAGTAAGTATAGTACCAAGAGAA 1943
QY 781 TCCACACTGTGTGCTCTTTTACAAAGCTATCAAGTCTGGAAGATGACCCAACTTCCAAAAC 840
Db 1944 TCCACACTGTGTGCTCTTTTACAAAGCTATCAAGTCTGGAAGATGACCCAACTTCCAAAAC 2003
QY 841 AAAGGAATTTTGGGCCACCAATTTTGGATTTCAGAGGTCTCTTCTGACTATATCTCTCC 900
Db 2004 AAAGGAATTTTGGGCCACCAATTTTGGATTTCAGAGGTCTCTTCTGACTATATCTCTCC 2063
QY 901 AGAAATTTTCTCTTTTGAAGTACTACTGATTTTACATTTATGATGGATGCTCTACAAGCC 960
Db 2064 AGAAATTTTCTCTTTTGAAGTACTACTGATTTTACATTTATGATGGATGCTCTACAAGCC 2123
QY 961 TCATGATCTACAGCTGGAAAGAAATATCTTACTGTGCTGTTTCATATATGGTGGTC --- 1016
Db 2124 TCATGATCTACAGCTGGAAAGAAATATCTTACTGTGCTGTTTCATATATGGTGGTC 2183
QY 1017 ----- 1016
Db 2184 CAATAGAAATTCAGCATCAGTGGGAAGGACTCCAATATCTAGCTTCTCGATATGATTC 2243
QY 1017 ----- 1016
Db 2244 ATTGACTTAGATCGTGTGGGCATCCAGGCTGGTCTATGAGAGGATACCTCTCCTCGATG 2303
QY 1017 -----CTCAGGTTGCTATTGCTGGGGCCCCAGTCACTCTG 1051
Db 2304 GCATTAATGCAGAGTCAGATATCTTCAGGGTTGCTATTGCTGGGGCCCCAGTCACTCTG 2363
QY 1052 TGGATCTTATGATACAGGATACAGGAAAGTTATATGGGTACCCCTGACCAAGATGAA 1111
Db 2364 TGGATCTTATGATACAGGATACAGGAAAGTTATATGGGTACCCCTGACCAAGATGAA 2423
QY 1112 CAGGCTATTACTTATGATCTGGCCATGCAAGCAGAAAGTTCCCTCTGACCAAAAT 1171
Db 2424 CAGGCTATTACTTATGATCTGGCCATGCAAGCAGAAAGTTCCCTCTGACCAAAAT 2483
QY 1172 CGTTTACTGCTTTACATGGTTTCTTGGATGAGAATGCCATTTTGCACATACCAGTATA 1231
Db 2484 CGTTTACTGCTTTACATGGTTTCTTGGATGAGAATGCCATTTTGCACATACCAGTATA 2543
QY 1232 TTACTAGTTTTTTAGTAGGGCTGGAAGCCATATGATTTTACAGATCTATCTCAGGAG 1291
Db 2544 TTACTAGTTTTTTAGTAGGGCTGGAAGCCATATGATTTTACAGATCTATCTCAGGAG 2603
QY 1292 AGACACAGCATAGAGTTCTGAACTCGGAGAACATATGAAGTGCATCTTTTGCATAC 1351
Db 2604 AGACACAGCATAGAGTTCTGAACTCGGAGAACATATGAAGTGCATCTTTTGCATAC 2663
QY 1352 CTTTCAAGAAACCTTGGATCAGCTATTGCTGCTTAAAGTGCATATAATTTTGAACCTGTG 1411
Db 2664 CTTTCAAGAAACCTTGGATCAGCTATTGCTGCTTAAAGTGCATATAATTTTGAACCTGTG 2723
QY 1412 TAGAACTCTCTGGTATACACTGGCTATTTTAAACAAATGAGGAGGTTTAAATCAACAGAAA 1471

Db 2724 TAGAACTCTCTGGTATACACTGGCTATTTTAAACAAATGAGGAGGTTTAAATCAACAGAAA 2783
QY 1472 CACAGAAATTTGATCATCATCATTTTGTATACCTGCGCATGTAACATCTACTCTCTGAAAAATAAT 1531
Db 2784 CACAGAAATTTGATCATCATCATTTTGTATACCTGCGCATGTAACATCTACTCTCTGAAAAATAAT 2843
QY 1532 GTGGTGCCATGCGAGGGTCTACCGTTTGTGGTAGTAAATCTAATACTTTAAACCCCATGTC 1591
Db 2844 GTGGTGCCATGCGAGGGTCTACCGTTTGTGGTAGTAAATCTAATACTTTAAACCCCATGTC 2903
QY 1592 TCAAAATCAATGATACATATTCCTGAGAGACCCAGCAATACCATAGAAATTTACTAAAAA 1651
Db 2904 TCAAAATCAATGATACATATTCCTGAGAGACCCAGCAATACCATAGAAATTTACTAAAAA 2963
QY 1652 AAAAAAAAAAAAAAAAAA 1668
Db 2964 AAAAAAAAAAAAAAAAAA 2980

RESULT 7
US-10-170-789-37
; Sequence 37, Application US/10170789
; Publication No. US20030180930A1
; GENERAL INFORMATION:
; APPLICANT: Rachel E. Meyers
; APPLICANT: Olandt, Peter J.
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Curtis, Rory A. J.
; APPLICANT: Williamson, Mark
; APPLICANT: Welch, Nadine
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASE, PHOSPHATASE,
; FILE OF INVENTION: AND PROTEASE FAMILY MEMBERS AND USES THEREOF
; FILE REFERENCE: 10448-191001
; CURRENT APPLICATION NUMBER: US/10/170, 789
; CURRENT FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: US 09/797,039
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: PCT/US01/06525
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/186,061
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 09/882,166
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: PCT/US01/19269
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/212,078
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 09/934,406
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: PCT/US01/26052
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/226,740
; PRIOR FILING DATE: 2000-08-21
; PRIOR APPLICATION NUMBER: US 09/861,801
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: PCT/US01/16549
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: US 60/205,508
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: US 09/801,267
; PRIOR FILING DATE: 2001-03-06
; PRIOR APPLICATION NUMBER: PCT/US01/07138
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: US 60/187,454
; PRIOR FILING DATE: 2000-03-07
; PRIOR APPLICATION NUMBER: US 09/829,671
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: PCT/US01/40483
; PRIOR FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: US 60/197,508
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: US 09/961,721

		PRIOR FILING DATE: 2001-09-24		PCT/US01/29904	
		PRIOR FILING DATE: 2001-09-24		US 60/235,023	
		PRIOR FILING DATE: 2000-09-25		US 10/045,367	
		PRIOR FILING DATE: 2001-11-07		US 60/246,561	
		PRIOR FILING DATE: 2000-11-07		US 09/801,275	
		PRIOR FILING DATE: 2001-03-06		PCT/US01/07074	
		PRIOR FILING DATE: 2001-03-05		US 60/187,420	
		PRIOR FILING DATE: 2000-03-07		NUMBER OF SEQ ID NOS: 63	
		SOFTWARE: Fast-Seq for Windows Version 4.0			
		SEQ ID NO 37			
		LENGTH: 3143			
		TYPE: DNA			
		ORGANISM: Homo sapiens			
		FEATURE:			
		NAME/KEY: CDS			
		LOCATION: (229)...(2874)			
		US-10-170-789-37			
		Query Match		81.1%; Score 1353.4; DB 6; Length 3143;	
		Best Local Similarity		84.7%; Pred. No. 0;	
		Matches 1664; Conservative		0; Mismatches 1; Indels 300; Gaps 1;	
Qy	1	AACAGGTACAGCAAAATCCCTAAAGTCACTTTTAAAGATGTCAGAAATAATGATGATGCTGA	60		
Db	1179	AACAGGTACAGCAAAATCCCTAAAGTCACTTTTAAAGATGTCAGAAATAATGATGATGCTGA	1238		
Qy	61	AGGAAGATCATAGATGTCATAGATAAGGAACATAATCAACCTTTTGAGATTCCTATTGGA	120		
Db	1239	AGGAAGATCATAGATGTCATAGATAAGGAACATAATCAACCTTTTGAGATTCCTATTGGA	1298		
Qy	121	AGGAGTTGAATATATTGCCAGAGCTGGATGGAATCCTCAGGGAATAATGCTTGGTCCAT	180		
Db	1299	AGGAGTTGAATATATTGCCAGAGCTGGATGGAATCCTCAGGGAATAATGCTTGGTCCAT	1358		
Qy	181	CCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCAGCTCACTGAATATTATAT	240		
Db	1359	CCTACTAGATCGCTCCAGACTCGCTCAGATAGTGTGATCTCAGCTCACTGAATATTATAT	1418		
Qy	241	CCAGTAGAAGATGATGTTATGGAAGGCAGAGACTCATTGAGTCAGTGGCTGATTCGT	300		
Db	1419	CCAGTAGAAGATGATGTTATGGAAGGCAGAGACTCATTGAGTCAGTGGCTGATTCGT	1478		
Qy	301	GACGCCACTAATTTATCTATGAGAACACACAGACATCTGGATAAATATCCATGACATCTT	360		
Db	1479	GACGCCACTAATTTATCTATGAGAACACACAGACATCTGGATAAATATCCATGACATCTT	1538		
Qy	361	TCATGTTTTTCCCAAGTCACGAGAGGAATTCAGTTTATTTTTCGCTCTGAAAGCA	420		
Db	1539	TCATGTTTTTCCCAAGTCACGAGAGGAATTCAGTTTATTTTTCGCTCTGAAAGCA	1598		
Qy	421	AACAGGTTCCGTCATTATATACAAATTTACATCTATTTTAAAGGAAGCAATATAACG	480		
Db	1599	AACAGGTTCCGTCATTATATACAAATTTACATCTATTTTAAAGGAAGCAATATAACG	1658		
Qy	481	ATCCAGTGGTGGCTGCCTCTCCCAAGTATTTCAAGTGTCTATCAAGAGGATAGC	540		
Db	1659	ATCCAGTGGTGGCTGCCTCTCCCAAGTATTTCAAGTGTCTATCAAGAGGATAGC	1718		
Qy	541	AATTAACAGTGGTGAATGGGAAGTTCTTGGCCGGCATGGATCTAATATCCAAGTTGATGA	600		
Db	1719	AATTAACAGTGGTGAATGGGAAGTTCTTGGCCGGCATGGATCTAATATCCAAGTTGATGA	1778		
Qy	601	AGTCAGAGGCTGGTATATTTTGAAGGCCACCAAGACTCCCTTTTAGAGCATCACCTGTA	660		
Db	1779	AGTCAGAGGCTGGTATATTTTGAAGGCCACCAAGACTCCCTTTTAGAGCATCACCTGTA	1838		
Qy	661	CGTAGTCAGTTACGTAAATCCCTGGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACA	720		
Db	1839	CGTAGTCAGTTACGTAAATCCCTGGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACA	1898		
Qy	721	TTCTTGTCTGCATCAGTCAGCAGCTGTGACTTCTTTTAAAGTAAAGTATAGTAACCAAGAA	780		
Db	1899	TTCTTGTCTGCATCAGTCAGCAGCTGTGACTTCTTTTAAAGTAAAGTATAGTAACCAAGAA	1958		
Qy	781	TCCACACTGTGTCTCCCTTTTACAAGCTATCAAGTCTCTGAAGATGACCCAACTTCGAAAAC	840		
Db	1959	TCCACACTGTGTCTCCCTTTTACAAGCTATCAAGTCTCTGAAGATGACCCAACTTCGAAAAC	2018		
Qy	841	AAAGGAATTTTGGGCCACCACTTTTGGATTTCAGCAGGTCTCTCTTCCTGACTATATCTCTCC	900		
Db	2019	AAAGGAATTTTGGGCCACCACTTTTGGATTTCAGCAGGTCTCTCTTCCTGACTATATCTCTCC	2078		
Qy	901	AGAAATTTTCTCTTTTGAAGTACTACTGGAATTTACATTTGATGGGATGCTCTCAAGGCC	960		
Db	2079	AGAAATTTTCTCTTTTGAAGTACTACTGGAATTTACATTTGATGGGATGCTCTCAAGGCC	2138		
Qy	961	TCATGATCTACAGCCTGGAAAGAAATATCTTACTGTCTGTCTTTCATATATGTTGGTCTCTCA	1020		
Db	2139	TCATGATCTACAGCCTGGAAAGAAATATCTTACTGTCTGTCTTTCATATATGTTGGTCTCTCA	2198		
Qy	1021	-----	-----	-----	-----
Db	2199	GGTGCAGTTGGTGAATAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTTAGC	2258		
Qy	1021	-----	-----	-----	-----
Db	2259	CTCTTAGGTTATGTGGTTGTAGTAGACAACAGGGGATCCCTGTACCAGGGCTTAA	2318		
Qy	1021	-----	-----	-----	-----
Db	2319	ATTTGAAGCGCCTTTAAATATAAATATGAGTCAATAGAAATTCAGATCAGGTGGGAAGG	2378		
Qy	1021	-----	-----	-----	-----
Db	2379	ACTCAATATCTAGCTTCTCGATATAGATTTCATTCGATCGTGGGATCCACGG	2438		
Qy	1021	-----	-----	-----	-----
Db	2439	CTGGTCTCTATGGAGGATACCTCTCCCTGATGGCATTAAATGCAGAGGTGAGATATCTTCAG	2498		
Qy	1021	GGTTGCTATTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGA	1080		
Db	2499	GGTTGCTATTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGA	2558		
Qy	1081	ACGTTATATGGGTCAACCTGACCCAGAAATGAAAGGGCTATTACTTAGGATCTGTGGCCAT	1140		
Db	2559	ACGTTATATGGGTCAACCTGACCCAGAAATGAAAGGGCTATTACTTAGGATCTGTGGCCAT	2618		
Qy	1141	GCAAGCAGAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGTTTCTCTGGA	1200		
Db	2619	GCAAGCAGAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGTTTCTCTGGA	2678		
Qy	1201	TGGAATGTCATTTTGGCAGTACAGTATATTACTGAGTTTCTTTTGTAGGGGCTGAAA	1260		
Db	2679	TGGAATGTCATTTTGGCAGTACAGTATATTACTGAGTTTCTTTTGTAGGGGCTGAAA	2738		
Qy	1261	GCCATATGATTTACAGATCTATCTCAGAGAGACAGACATAGAGTCTGATCGGG	1320		
Db	2739	GCCATATGATTTACAGATCTATCTCAGAGAGACAGACATAGAGTCTGATCGGG	2798		
Qy	1321	AGAACATTTAGACTGCTTTTGGCACTACCTTCAAGAAACCTTGGATCAGGTATTGC	1380		
Db	2799	AGAACATTTAGACTGCTTTTGGCACTACCTTCAAGAAACCTTGGATCAGGTATTGC	2858		
Qy	1381	TGCTCTAAAAAGTGAATAAATTTGACCTGTGTAGAACTCTCTGGTATACCTGGCTATTT	1440		
Db	2859	TGCTCTAAAAAGTGAATAAATTTGACCTGTGTAGAACTCTCTGGTATACCTGGCTATTT	2918		

QY 1441 AACCAATGAGGAGTTTAATCAACAGAAACAGAAATGATCATCATCTTTTGATACC 1500
DB 2919 AACCAATGAGGAGTTTAATCAACAGAAACAGAAATGATCATCATCTTTTGATACC 2978
QY 1501 TGCCATGTAAACATCTACTCTCTGAAATAAATGTTGGTCCCATGAGGGGCTACGGTTTGT 1560
DB 2979 TGCCATGTAAACATCTACTCTCTGAAATAAATGTTGGTCCCATGAGGGGCTACGGTTTGT 3038
QY 1561 GGTAGTAATCTAATACCTTAACCCACATGCTCAAAATCAATGATACATATTTCTGTAGA 1620
DB 3039 GGTAGTAATCTAATACCTTAACCCACATGCTCAAAATCAATGATACATATTTCTGTAGA 3098
QY 1621 GACCAGCAATACCAATGAGAAATTAATAAAAAAAAAAAAAAAAAAAAA 1665
DB 3099 GACCAGCAATACCAATGAGAAATTAATAAAAAAAAAAAAAAAAAAAAA 3143

RESULT 8
US-09-976-674-12
; Sequence 12, Application US/09976674
; Patent No. US20020115843A1
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 4829
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-976-674-12

Query Match 80.9%; Score 1349.6; DB 3; Length 4829;
Best Local Similarity 84.5%; Pred. No. 0;
Matches 1664; Conservative 0; Mismatches 4; Indels 302; Gaps 1;

QY 1 AACAGGTACAGCAATCTTAAAGTCACTTTTAAAGATGTCAGAAATATGATGCTGA 60
DB 1164 AACAGGTACAGCAATCTTAAAGTCACTTTTAAAGATGTCAGAAATATGATGCTGA 1223
QY 61 AGGAAGATCATAGATGTCATAGATAAGAACTAAATCAACCTTTTGAGATTCTATTGA 120
DB 1224 AGGAAGATCATAGATGTCATAGATAAGAACTAAATCAACCTTTTGAGATTCTATTGA 1283
QY 121 AGGATTTGATATATTTGCCAGAGCTGGATGACTCTCGAGGAAATATGCTGGTCCAT 180
DB 1284 AGGATTTGATATATTTGCCAGAGCTGGATGACTCTCGAGGAAATATGCTGGTCCAT 1343
QY 181 CTTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACTCACTGAATTTAT 240
DB 1344 CTTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACTGAATTTAT 1403
QY 241 CCCAGTAGAAGATGATGTTATGGAAGGAGAGAGCTCAATGAGTCAGTGGCTGATTCGT 300
DB 1404 CCCAGTAGAAGATGATGTTATGGAAGGAGAGAGCTCAATGAGTCAGTGGCTGATTCGT 1463
QY 301 GAGCCCAATTAATCTATGAGAAACACAGACATCTGGATAAATATCCATGACATCTT 360
DB 1464 GAGCCCAATTAATCTATGAGAAACACAGACATCTGGATAAATATCCATGACATCTT 1523
QY 361 TCATGTTTTCCCAAGTCAGAGAGAAATTCAGTTTATTTTTCCTCTGAATGCAA 420
DB 1524 TCATGTTTTCCCAAGTCAGAGAGAAATTCAGTTTATTTTTCCTCTGAATGCAA 1583
QY 421 AACAGGTTTCCGTCATTATACAAATTACATCTATTTTAAAGGAAACCAATATAACG 480

DB 1584 AACAGGTTTCCGTCATTATATACAAATTTACATCTATTTTAAAGGAAACCAATATAACG 1643
QY 481 ATCCAGTGGTGGCTGCTGCTCCAGTGTATTTCAAGTGTCCCTATCAAGAGGAGATAGC 540
DB 1644 ATCCAGTGGTGGCTGCTGCTCCAGTGTATTTCAAGTGTCCCTATCAAGAGGAGATAGC 1703
QY 541 AATTACCAAGTGTGAATGGGAAGTTCTTGGCCGCATGGATCTAATATCAACGTTGATGA 600
DB 1704 AATTACCAAGTGTGAATGGGAAGTTCTTGGCCGCATGGATCTAATATCAACGTTGATGA 1763
QY 601 AGTCAGAAGGCTGTATATTTTGAAGGCACCAAGAGCTCCCTTTAGAGCATCACTGTGA 660
DB 1764 AGTCAGAAGGCTGTATATTTTGAAGGCACCAAGAGCTCCCTTTAGAGCATCACTGTGA 1823
QY 661 CGTAGTCAGTTACGTAATTCCTGGAGAGGTGACAAGGCTGACTGACCGTGGCTACTACA 720
DB 1824 CGTAGTCAGTTACGTAATTCCTGGAGAGGTGACAAGGCTGACTGACCGTGGCTACTACA 1883
QY 721 TTCTTGTCTCATCAGTCAGCACTGTGACTTCTTTATAGTAAGTATAGTAACCAAGAA 780
DB 1884 TTCTTGTCTCATCAGTCAGCACTGTGACTTCTTTATAGTAAGTATAGTAACCAAGAA 1943
QY 781 TCCACACTGTGTGCTCCCTTTTCAAGCTATCAAGTCTTGAAGATGACCCAACTTGCAGAAC 840
DB 1944 TCCACACTGTGTGCTCCCTTTTCAAGCTATCAAGTCTTGAAGATGACCCAACTTGCAGAAC 2003
QY 841 AAAGGAATTTTGGGCCACCAATTTTGGATTACAGAGGCTCTCTCTCTGACTATATCTCTCC 900
DB 2004 AAAGGAATTTTGGGCCACCAATTTTGGATTACAGAGGCTCTCTCTCTGACTATATCTCTCC 2063
QY 901 AGAAATTTTCTCTTTTGAAGTACTACTGGAATTTACATGTATGGATGCTCTACAAGCC 960
DB 2064 AGAAATTTTCTCTTTTGAAGTACTACTGGAATTTACATGTATGGATGCTCTACAAGCC 2123
QY 961 TCATGATCTACAGCTGGAGAAATATCTACTGCTGCTGTTTCATATATGTTGCTGCTCTCA 1020
DB 2124 TCATGATCTACAGCTGGAGAAATATCTACTGCTGCTGTTTCATATATGTTGCTGCTCTCT 2183
QY 1021 ----- 1020
DB 2184 CAGGTGAGTTGGTGAATTAATCGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCCTA 2243
QY 1021 ----- 1020
DB 2244 GCCTCTCTAGTTATCTGTTGTTAGTATAGACACAGGGGATCCTGTCCCGAGGGCTT 2303
QY 1021 ----- 1020
DB 2304 AAATTTGAAGGCGCTTTAAATATAAAATGGGTCAATAGAAATGACGATCAGGTGGAA 2363
QY 1021 ----- 1020
DB 2364 GGACTCCAATATCTAGCTTCTCGATATGATTTTCAATTGACTTAGATCGTGGGATCCAC 2423
QY 1021 ----- 1020
DB 2424 GGCTGTCTTATGAGAGATACCTCTCCCTGATGGCAATTAATGACAGAGTCAAGATATCTTC 2483
QY 1021 --GTTTCTATTTCTGGGGCCCGAGTCACTCTGTGGATCTTCTATGATACAGGATACACG 1078
DB 2484 AGGTTTCTATTTCTGGGGCCCGAGTCACTCTGTGGATCTTCTATGATACAGGATACACG 2543
QY 1079 GAAAGTTTATAGGTCACCTCTGACCAAGATGAAGAGGCTATTACTTAGGATCTGTGGCC 1138
DB 2544 GAAAGTTTATAGGTCACCTCTGACCAAGATGAAGAGGCTATTACTTAGGATCTGTGGCC 2603
QY 1139 ATCAAGCAAGAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGTTTCTCTG 1198
DB 2604 ATCAAGCAAGAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGTTTCTCTG 2663
QY 1199 GATGAGATGTCATTTTGGACATACAGTATATCTAGTGTGTTTATTTAGTGAGGCTGGA 1258

```
Db 2664 GATGAGAATGTCCATTTTGCACATACCAGTATATTACTGAGTTTTTTAGTGAGGGCTGGA 2723
Qy 1259 AAGCCATATGATTTACAGATCTATCTCAGGAGAGACACAGCATAGAGTTCTTGATCG 1318
Db 2724 AAGCCATATGATTTACAGATCTATCTCAGGAGAGACACAGCATAGAGTTCTTGATCG 2783
Qy 1319 GGAGACATATGAGCTGATCTTTGGACCTACCTTTCAAGAAACCTTGGATCAGTATT 1378
Db 2784 GGAGACATATGAGCTGATCTTTGGACCTACCTTTCAAGAAACCTTGGATCAGTATT 2843
Qy 1379 GCTGCTTAAAGTGATTAATTTGACCTGTAGAACTCTCTGGTATACCTGGCTAT 1438
Db 2844 GCTGCTTAAAGTGATTAATTTGACCTGTAGAACTCTCTGGTATACCTGGCTAT 2903
Qy 1439 TTAACCAATGAGGAGGTTTAAATCAACAGAAACACAGAAATGTATCATCATTTTGATA 1498
Db 2904 TTAACCAATGAGGAGGTTTAAATCAACAGAAACACAGAAATGTATCATCATTTTGATA 2963
Qy 1499 CTGCGATGTAACTACTCTCTGAAAATAAATGTGTGCCATGCAGGGGTCTACGGTTT 1558
Db 2964 CTGCGATGTAACTACTCTCTGAAAATAAATGTGTGCCATGCAGGGGTCTACGGTTT 3023
Qy 1559 GTGGTAGTAATCTAATACCTTAACCCACATGCTCAAAATCAAAATGATATATCTCTGA 1618
Db 3024 GTGGTAGTAATCTAATACCTTAACCCACATGCTCAAAATCAAAATGATATATCTCTGA 3083
Qy 1619 GAGACCCAGCAATACCAATAGAAATTAACCAAGAAATTAACCAAGAAATTAACCAAGAA 1668
Db 3084 GAGACCCAGCAATACCAATAGAAATTAACCAAGAAATTAACCAAGAAATTAACCAAGAA 3133

RESULT 9
US-10-982-512-12
; Sequence 12, Application US/10982512
; Publication No. US20050059081A1
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/10/982,512
; PRIOR FILING DATE: 2004-11-05
; PRIOR FILING DATE: US/09/976,674
; PRIOR FILING DATE: 2001-10-12
; PRIOR FILING DATE: US 60/240,117
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 4829
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-982-512-12

Query Match 80.9%; Score 1349,6; DB 9; Length 4829;
Best Local Similarity 84.5%; Pred. No. 0;
Matches 1664; Conservative 0; Mismatches 4; Indels 302; Gaps 1;

Qy 1 AACAGGTACAGCAATCTTAAGTCACCTTTTAAAGTGTACAGAAATTAAGTATGATCGTGA 60
Db 1164 AACAGGTACAGCAATCTTAAGTCACCTTTTAAAGTGTACAGAAATTAAGTATGATCGTGA 1223
Qy 61 AGGAGGATCATAGTGTATAGATAGAGCACTTAATCAACCTTTGAGATTCATTTTGA 120
Db 1224 AGGAGGATCATAGTGTATAGATAGAGCACTTAATCAACCTTTGAGATTCATTTTGA 1283
Qy 121 AGGAGTGAATATATTTGCCAGAGCTGGATGACTCTCTGAGGGAAATATGCTTGGTCCAT 180
Db 1284 AGGAGTGAATATATTTGCCAGAGCTGGATGACTCTCTGAGGGAAATATGCTTGGTCCAT 1343
Qy 181 CCTACTAGATCGCTCCAGACTCGCCTACAGATAGTGTGATCTCACCTGAATATTATTTAT 240
```

```
Db 1344 CCTACTAGATCGCTCCAGACTCGCCTACAGATAGTGTGTATCTCACCTGAATATTAT 1403
Qy 241 CCCAGTAAAGATGATGTTATGGAAAGGAGAGACTCATTTGAGTCAGTGCCTGATTTCTGT 300
Db 1404 CCCAGTAAAGATGATGTTATGGAAAGGAGAGACTCATTTGAGTCAGTGCCTGATTTCTGT 1463
Qy 301 GAGCCCACTAATTTATCTATGAAGAAACAAACAGACATCTGGATAAATATCCATGACATCTT 360
Db 1464 GAGCCCACTAATTTATCTATGAAGAAACAAACAGACATCTGGATAAATATCCATGACATCTT 1523
Qy 361 TCATGTTTTTCCCAAGTCAAGAGAGGAAATTCAGTTTATTTTTTGGCTCTGTAATGCAA 420
Db 1524 TCATGTTTTTCCCAAGTCAAGAGAGGAAATTCAGTTTATTTTTTGGCTCTGTAATGCAA 1583
Qy 421 AACAGGTTTCCGTCATTTATACAAATATACATCTATTTTAAAGGAAACCAATATAAAG 480
Db 1584 AACAGGTTTCCGTCATTTATACAAATATACATCTATTTTAAAGGAAACCAATATAAAG 1643
Qy 481 ATCCAGTGTGGGCTGCCTCCCAAGTGATTTCAAGTGTCTCTATCAAGAGGAGATAGC 540
Db 1644 ATCCAGTGTGGGCTGCCTCCCAAGTGATTTCAAGTGTCTCTATCAAGAGGAGATAGC 1703
Qy 541 AATTACAGTGTGATGATGGAGTTTCTTGGCCGATGGATCTAATATCCAAAGTTGATGA 600
Db 1704 AATTACAGTGTGATGATGGAGTTTCTTGGCCGATGGATCTAATATCCAAAGTTGATGA 1763
Qy 601 AGTCAGAGGCTGGTATATTTTGAAGGCCAACAAAGACTCCCTTTTAGAGATCACTCTGTA 660
Db 1764 AGTCAGAGGCTGGTATATTTTGAAGGCCAACAAAGACTCCCTTTTAGAGATCACTCTGTA 1823
Qy 661 CGTAGTCAGTTACGTAATTCCTGGAGAGGTGACAAAGGCTGACTGACCGTGGTACTCACA 720
Db 1824 CGTAGTCAGTTACGTAATTCCTGGAGAGGTGACAAAGGCTGACTGACCGTGGTACTCACA 1883
Qy 721 TTCTTGCTGCATCAGTCAGCACTGTGACTTCTTTATAAGTAAGTATAGTAAACAGAGAA 780
Db 1884 TTCTTGCTGCATCAGTCAGCACTGTGACTTCTTTATAAGTAAGTATAGTAAACAGAGAA 1943
Qy 781 TCACACTGTGTGCTTCTTTGAAAGTACTACTGATTTTACATTTGATGGGATGCTCTACAGCC 840
Db 1944 TCACACTGTGTGCTTCTTTGAAAGTACTACTGATTTTACATTTGATGGGATGCTCTACAGCC 2003
Qy 841 AAAGGAAATTTGGGCCACCATTTTGGATTTCAGCAGGTCTCTTCTGACTATATCTCTCC 900
Db 2004 AAAGGAAATTTGGGCCACCATTTTGGATTTCAGCAGGTCTCTTCTGACTATATCTCTCC 2063
Qy 901 AGAAATTTTCTCTTTTGAAGTACTACTGATTTTACATTTGATGGGATGCTCTACAGCC 960
Db 2064 AGAAATTTTCTCTTTTGAAGTACTACTGATTTTACATTTGATGGGATGCTCTACAGCC 2123
Qy 961 TCATGATCTACAGCTGAGAAATATCCTGCTGCTGCTCATATATATGGTGGTCTCA 1020
Db 2124 TCATGATCTACAGCTGAGAAATATCCTGCTGCTGCTCATATATATGGTGGTCTCCT 2183
Qy 1021 ----- 1020
Db 2184 CAGGTGCGAGTGGTGATTAATTCGGTTTAAAGGAGTCAAGTATTTCCGCTTTGAATACCTTA 2243
Qy 1021 ----- 1020
Db 2244 GCCTCTCAGGTTATGTTGGTTTAGTGAATAGACAAAGGGGATCTGTCTACCCAGGGCTT 2303
Qy 1021 ----- 1020
Db 2304 AAATTTGAAGGGCTTTTAAATATATAAATGGGTCAAAATAGAAATTTGACGATCAGGTGGAA 2363
Qy 1021 ----- 1020
Db 2364 GGACTCCAAATATCTAGCTTCTCGATATGATTTTCATTGACTTAGATCGTGTGGGCGATCCAC 2423
Qy 1021 ----- 1020
```

Db 2424 GGCTGTCCTATGAGGATACCTCTCCCTGATGGCAATTAATGCGAGGTCAGATATCTTC 2483
Qy 1021 --GGTGTCTATTCTCGGGCCCGAGTCACTCTGTGGATCTTCTATGATACAGGATACACG 1078
Db 2484 AGGTTGCTATTCTCGGGCCCGAGTCACTCTGTGGATCTTCTATGATACAGGATACACG 2543
Qy 1079 GAACGTTATATGGGTCACTCCCTGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCC 1138
Db 2544 GAACGTTATATGGGTCACTCCCTGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCC 2603
Qy 1139 ATGCAAGCAGAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGTTTCCCTG 1198
Db 2604 ATGCAAGCAGAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGTTTCCCTG 2663
Qy 1199 GATGAGAAATGTCATTTTGGACATACCAATGATATATTACTGAGTTTATGAGGGCTGGA 1258
Db 2664 GATGAGAAATGTCATTTTGGACATACCAATGATATATTACTGAGTTTATGAGGGCTGGA 2723
Qy 1259 AAGCCATATGATTTTACAGATCTATCTCAGGAGACACAGCATAAGAGTTTCTCGAATCG 1318
Db 2724 AAGCCATATGATTTTACAGATCTATCTCAGGAGACACAGCATAAGAGTTTCTCGAATCG 2783
Qy 1319 GGAGCAATATGAACTGCACTCTTTGCACTACCTTCAAGAAAACCTTGGATCACTGATTT 1378
Db 2784 GGAGCAATATGAACTGCACTCTTTGCACTACCTTCAAGAAAACCTTGGATCACTGATTT 2843
Qy 1379 GCTGCTCTAAAAGTGATATAATTTTGACCTGTGTAGAACTCTCTGGTATACACTGGCTAT 1438
Db 2844 GCTGCTCTAAAAGTGATATAATTTTGACCTGTGTAGAACTCTCTGGTATACACTGGCTAT 2903
Qy 1439 TTAAACCAATGAGGAGTTTAAATCAACAGAAAACACAGAAATGATATCAATTTGATA 1498
Db 2904 TTAAACCAATGAGGAGTTTAAATCAACAGAAAACACAGAAATGATATCAATTTGATA 2963
Qy 1499 CTTGCCATGTAATCTCTCTGAAAATAAATGTTGGTGCATGAGGGGTCTACGGTTT 1558
Db 2964 CTTGCCATGTAATCTCTCTGAAAATAAATGTTGGTGCATGAGGGGTCTACGGTTT 3023
Qy 1559 GTGGTAGTAATCTAAATACCTTAACCCCATGCTCAAAATCAAAATGATATATTTCTGA 1618
Db 3024 GTGGTAGTAATCTAAATACCTTAACCCCATGCTCAAAATCAAAATGATATATTTCTGA 3083
Qy 1619 GAGACCCAGCAATACCATAGAAATTAATAAAAAAATAAAAAAATAAAAAA 1668
Db 3084 GAGACCCAGCAATACCATAGAAATTAATAAAAAAATAAAAAAATAAAAAA 3133

RESULT 10

US-10-415-122-5
; Sequence 5, Application US/10415122
; Publication No. US20040053369A1
; GENERAL INFORMATION:
; APPLICANT: THE UNIVERSITY OF SYDNEY
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES
; FILE REFERENCE: FFI5217
; CURRENT APPLICATION NUMBER: US/10/415,122
; CURRENT FILING DATE: 2003-08-07
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 3120
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-415-122-5

Query Match

Best Local Similarity 84.7%; Score 1347; DB 7; Length 3120;
Matches 1657; Conservative 0; Mismatches 0; Indels 300; Gaps 1;
Qy 1 AACAGGTACAGCAAAATCCTTAAGTCACTTTTAAGATGTCAGAAATATATGATGCTGA 60
Db 1164 AACAGGTACAGCAAAATCCTTAAGTCACTTTTAAGATGTCAGAAATATATGATGCTGA 1223

Qy 61 AGGAAGGATCATAGATGTCTATAGATAAGAACTAAATTCACCTTTTGAGATTCTATTGA 120
Db 1224 AGGAAGGATCATAGATGTCTATAGATAAGAACTAAATTCACCTTTTGAGATTCTATTGA 1283
Qy 121 AGGAGTTGAATATATTGCGAGAGCTGGATGGAATCTCTGAGGAAATATGTTGTCAT 180
Db 1284 AGGAGTTGAATATATTGCGAGAGCTGGATGGAATCTCTGAGGAAATATGTTGTCAT 1343
Qy 181 CCTACTAGATCGCTCCAGAGCTCGCTACAGATAGTGTGTGATCTCACCTGAATATTAT 240
Db 1344 CCTACTAGATCGCTCCAGAGCTCGCTACAGATAGTGTGTGATCTCACCTGAATATTAT 1403
Qy 241 CCCAGTAGAAGATGATGTTTATGAAAGGACAGAGACTCATTGAGTCAGTCGCTGATCTGT 300
Db 1404 CCCAGTAGAAGATGATGTTTATGAAAGGACAGAGACTCATTGAGTCAGTCGCTGATCTGT 1463
Qy 301 GACGCCACTAATTTCTATGAAGAAACAAACAGACATCTGGATAAATATCCATGACATCTT 360
Db 1464 GACGCCACTAATTTCTATGAAGAAACAAACAGACATCTGGATAAATATCCATGACATCTT 1523
Qy 361 TCATGTTTTTCCCAAGAGTCACGAGAGAAATTTGAGTTTATTTTGGCTCTGAATGCAA 420
Db 1524 TCATGTTTTTCCCAAGAGTCACGAGAGAAATTTGAGTTTATTTTGGCTCTGAATGCAA 1583
Qy 421 AACAGGTTTCGGTCATTTATACAAATTTACATCTATTTTAAAGGAAAGCAATATAAAG 480
Db 1584 AACAGGTTTCGGTCATTTATACAAATTTACATCTATTTTAAAGGAAAGCAATATAAAG 1643
Qy 481 ATCCAGTGGTGGGCTGCTCTCCAAAGTATTTCAAGTGTCTATCAAGAGGAGATAGC 540
Db 1644 ATCCAGTGGTGGGCTGCTCTCCAAAGTATTTCAAGTGTCTATCAAGAGGAGATAGC 1703
Qy 541 AATTACCAAGTGGTGAATGGGAAGTTCTTGCCGGCATGGATCTAATATCCAAAGTTGATGA 600
Db 1704 AATTACCAAGTGGTGAATGGGAAGTTCTTGCCGGCATGGATCTAATATCCAAAGTTGATGA 1763
Qy 601 AGTCAGAAGGCTGGTATATTTTGAAGGACCAAGACTCCCTTTAGAGCATCACCTGTA 660
Db 1764 AGTCAGAAGGCTGGTATATTTTGAAGGACCAAGACTCCCTTTAGAGCATCACCTGTA 1823
Qy 661 CGTAGTCAGTTACGTAATCTCTGAGAGGTGACAGGCTGACGACCGTGGTACTCACA 720
Db 1824 CGTAGTCAGTTACGTAATCTCTGAGAGGTGACAGGCTGACGACCGTGGTACTCACA 1883
Qy 721 TTTCTGTGTCATCAGTCAGCACTGTGACTTCTTTATAGTAGTAGTAGTAACAGAGAA 780
Db 1884 TTTCTGTGTCATCAGTCAGCACTGTGACTTCTTTATAGTAGTAGTAACAGAGAA 1943
Qy 781 TCCACACTGTGTGTCCTTTTACAAAGCTATCAAGTCTCTGAAGATGACCAACTTGCAAAAC 840
Db 1944 TCCACACTGTGTGTCCTTTTACAAAGCTATCAAGTCTCTGAAGATGACCAACTTGCAAAAC 2003
Qy 841 AAAGGAATTTTGGGCCACCACTTTTGGATTTCAGCAGGTCTCTTCTGACTATPACTCTCC 900
Db 2004 AAAGGAATTTTGGGCCACCACTTTTGGATTTCAGCAGGTCTCTTCTGACTATPACTCTCC 2063
Qy 901 AGAAATTTTCTCTTTTGAAGTACTACTGGAATTTACATTTGATGGATGCTCTACAAGCC 960
Db 2064 AGAAATTTTCTCTTTTGAAGTACTACTGGAATTTACATTTGATGGATGCTCTACAAGCC 2123
Qy 961 TCATGATCTACAGCTGGAAAGAAATATCTCTCTGTGTGTTTCAATATATGTTGGTGTCTCA 1020
Db 2124 TCATGATCTACAGCTGGAAAGAAATATCTCTCTGTGTGTTTCAATATATGTTGGTGTCTCA 2183
Qy 1021 ----- 1020
Db 2184 GGTGCACTGTGGTAATAATCGGTTTAAAGAGTCAAGTATTTTCGCTTGAATACCCCTAGC 2243
Qy 1021 ----- 1020
Db 2244 CTCTTAGGTTATGTGGTTGTAGTAGTAGCAACAGGGGATCTCTGTCCCGAGGGCTTAA 2303
Qy 1021 ----- 1020

Db 2304 ATTTGAAGGCGCTTTAAATATATAAAATGGGTCAAAATAGAAATTCAGCATCAGGTGGAGG 2363
Qy 1021 ----- 1020
Db 2364 ACTCCAATATCTAGCTTCGATATGATTTTCATTGACTTTAGATCGTGTGGGCATCCACGG 2423
Qy 1021 ----- 1020
Db 2424 CTGGTCTATGAGAGATACCTCTCCCTGATGGCATTAAATGCAGAGGTTCAGATATCTTCAG 2483
Qy 1021 GGTTCCTATTCCTGGGGCCCCAGTCACCTCTGTGGATCTTCTATGATACAGGATACACGGA 1080
Db 2484 GGTTCCTATTCCTGGGGCCCCAGTCACCTCTGTGGATCTTCTATGATACAGGATACACGGA 2543
Qy 1081 ACCTTATATGGGTTCACCTCTGACAGAAATGAACAGGGCTATTACTTAGGATCTGTGGCCAT 1140
Db 2544 ACCTTATATGGGTTCACCTCTGACAGAAATGAACAGGGCTATTACTTAGGATCTGTGGCCAT 2603
Qy 1141 GCAAGCAGAAAAGTTCCTCTGAAACCAATCGTTTACTCTCTTACATGTTTCTCTGGA 1200
Db 2604 GCAAGCAGAAAAGTTCCTCTGAAACCAATCGTTTACTCTCTTACATGTTTCTCTGGA 2663
Qy 1201 TGAGATGTCCTATTTGACATACCAATGATATTTACTGAGTTTCTTACTGAGGGCTGGAA 1260
Db 2664 TGAGATGTCCTATTTGACATACCAATGATATTTACTGAGTTTCTTACTGAGGGCTGGAA 2723
Qy 1261 GGCATATGATTTACAGATCTATCTCCAGGAGACACAGATAGAGTTCTCTGANTCGG 1320
Db 2724 GGCATATGATTTACAGATCTATCTCCAGGAGACACAGATAGAGTTCTCTGANTCGG 2783
Qy 1321 AGACATATGAACTGATCTTTTGGCATACCTTCAGAAAACCTTGGATCAGATATGC 1380
Db 2784 AGACATATGAACTGATCTTTTGGCATACCTTCAGAAAACCTTGGATCAGATATGC 2843
Qy 1381 TGCTCTAAAAGTGATATAATTTGACCTGTGTGAGAACTCTCTGTGATACACTGCTGATTT 1440
Db 2844 TGCTCTAAAAGTGATATAATTTGACCTGTGTGAGAACTCTCTGTGATACACTGCTGATTT 2903
Qy 1441 AACCAATGAGGAGGTTTAAATCAACAGAAAACACAGAAATGATCATCACTTTGTATAC 1500
Db 2904 AACCAATGAGGAGGTTTAAATCAACAGAAAACACAGAAATGATCATCACTTTGTATAC 2963
Qy 1501 TGCCATGTAACTACTCTCTGAAATAAATGTGTGTCATGCGGGGTCTACGGTTTGT 1560
Db 2964 TGCCATGTAACTACTCTCTGAAATAAATGTGTGTCATGCGGGGTCTACGGTTTGT 3023
Qy 1561 GGTAGTAATCTAATACCTTAAACCCACATGCTCAAAATCAAAATGATACATATTCCTGAGA 1620
Db 3024 GGTAGTAATCTAATACCTTAAACCCACATGCTCAAAATCAAAATGATACATATTCCTGAGA 3083
Qy 1621 GACCCAGCAATACCAATGAATTTACTTAAAAA 1657
Db 3084 GACCCAGCAATACCAATGAATTTACTTAAAAA 3120

RESULT 11
US-10-825-632-2
; Sequence 2 Application US/10825632
; Publication No US20040191826A1
; GENERAL INFORMATION:
; APPLICANT: ABBOTT; Catherine Anne
; TITLE OF INVENTION: DIFEPIDIL PEPTIDASES
; FILE REFERENCE: FCSB-100-Div. 1
; CURRENT APPLICATION NUMBER: US/10/825,632
; PRIOR FILING DATE: 2004-04-15
; PRIOR APPLICATION NUMBER: US 10/070,464
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: PCT/AU00/01085
; PRIOR FILING DATE: 2000-09-11
; PRIOR APPLICATION NUMBER: AU PQ5709
; PRIOR FILING DATE: 2000-02-18

; PRIOR APPLICATION NUMBER: AU PQ2762
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 3120
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-10-825-632-2

Query Match 80.7%; Score 1347; DB 8; Length 3120;
Best Local Similarity 84.7%; Pred. No. 0;
Matches 1657; Conservative 0; Mismatches 0; Indels 300; Gaps 1;
Qy 1 AACAGGTACAGCAAACTCTAAAGTCACCTTTAAAGATGTGAGAAATAAATGATGCTGA 60
Db 1164 AACAGGTACAGCAAACTCTAAAGTCACCTTTAAAGATGTGAGAAATAAATGATGCTGA 1223
Qy 61 AGGAAGGATCATAGATGTGATAGATAAGGAACATAATCAACCTTTGAGATTTCTATTGA 120
Db 1224 AGGAAGGATCATAGATGTGATAGATAAGGAACATAATCAACCTTTGAGATTTCTATTGA 1283
Qy 121 AGGAGTTGATATATTGCCAGAGCTGGATGGACTCTCGAGGAAATATGCTTGGTCCAT 180
Db 1284 AGGAGTTGATATATTGCCAGAGCTGGATGGACTCTCGAGGAAATATGCTTGGTCCAT 1343
Qy 181 CCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAATTTAT 240
Db 1344 CCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAATTTAT 1403
Qy 241 CCNAGTAGAGATGATTTATGGAAGGACAGACTCATTTGAGTCAGTGCCTGATTTCTGT 300
Db 1404 CCNAGTAGAGATGATTTATGGAAGGACAGACTCATTTGAGTCAGTGCCTGATTTCTGT 1463
Qy 301 GAGCCACTAATTTATCTATGAAAGAAACAAACAGACATCTGGATAAATATCCATGACATCTT 360
Db 1464 GAGCCACTAATTTATCTATGAAAGAAACAAACAGACATCTGGATAAATATCCATGACATCTT 1523
Qy 361 TCATGTTTTTCCCAAGTCACGAAAGGAAATTCAGTTTATTTTGGCTCTGAATGCAA 420
Db 1524 TCATGTTTTTCCCAAGTCACGAAAGGAAATTCAGTTTATTTTGGCTCTGAATGCAA 1583
Qy 421 AACAGGTTTCCTGCTCATTTATACAAAATTAATCATCTATTTTAAAGGAAAGCAAAATAAAGC 480
Db 1584 AACAGGTTTCCTGCTCATTTATACAAAATTAATCATCTATTTTAAAGGAAAGCAAAATAAAGC 1643
Qy 481 ATCCAGTGGTGGCTGCCTGCTCCAAAGTGATTTCAAGTGCTCTATCAAAAGAGGAGATAGC 540
Db 1644 ATCCAGTGGTGGCTGCCTGCTCCAAAGTGATTTCAAGTGCTCTATCAAAAGAGGAGATAGC 1703
Qy 541 AATTACCAGTGGTGAATGGAAAGTTCTTGGCCGGCATGGATCTAATATCCAAAGTTGATGA 600
Db 1704 AATTACCAGTGGTGAATGGAAAGTTCTTGGCCGGCATGGATCTAATATCCAAAGTTGATGA 1763
Qy 601 AGTCAGAAGGCTGGTATATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACCTGTA 660
Db 1764 AGTCAGAAGGCTGGTATATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACCTGTA 1823
Qy 661 GGTAGTCAGTACGTAATTCCTGGAGAGTGAGCAAGGCTGACTGACCGTGGCTACTCACA 720
Db 1824 GGTAGTCAGTACGTAATTCCTGGAGAGTGAGCAAGGCTGACTGACCGTGGCTACTCACA 1883
Qy 721 TTCTTCTGCATCAGTCAGCACTGTGCTCTTTTAAAGTAAGTATAGTACCAAGAGAA 780
Db 1884 TTCTTCTGCATCAGTCAGCACTGTGACTTTTAAAGTAAGTATAGTACCAAGAGAA 1943
Qy 781 TCNACTGTGTGCTCCCTTTACAGCTATCAAGTCTCAAGATGCCCACTTGCAGAAC 840
Db 1944 TCNACTGTGTGCTCCCTTTACAGCTATCAAGTCTCAAGATGCCCACTTGCAGAAC 2003
Qy 841 AAAGGAATTTTGGGCCACCATTTTGGATTTGAGATTTGAGAGGTCTCTTCTGATATCTCTCC 900
Db 2004 AAAGGAATTTTGGGCCACCATTTTGGATTTGAGATTTGAGAGGTCTCTTCTGATATCTCTCC 2063

Qy 301 GAGCCACTAATTATCTATGAAGAACACACAGACATCTGGATAATATCCATGACATCTT 360
Db 1453 GAGCCACTAATTATCTATGAAGAACACACAGACATCTGGATAATATCCATGACATCTT 1512
Qy 361 TCATGTTTTTCCCAAGTCACGAGAGGAAATGAGTTATTTTTTGGCTCTGAATGCAA 420
Db 1513 TCATGTTTTTCCCAAGTCACGAGAGGAAATGAGTTATTTTTTGGCTCTGAATGCAA 1572
Qy 421 AACAGGTTCCGTCATTTATACAAATTTACATCTATTTTAAAGGAAAGCAATATAAAG 480
Db 1573 AACAGGTTCCGTCATTTATACAAATTTACATCTATTTTAAAGGAAAGCAATATAAAG 1632
Qy 481 ATCCAGTGGGGCTGCCCTGCTCCAAAGTGATTTCAAGTGTCTTATCAAAAGAGGATAGC 540
Db 1633 ATCCAGTGGGGCTGCCCTGCTCCAAAGTGATTTCAAGTGTCTTATCAAAAGAGGATAGC 1692
Qy 541 AATTACAGTGGGAAATGGAAAGTTCTTGGCGGCATGGATCTAATATCCAAGTTGATGA 600
Db 1693 AATTACAGTGGGAAATGGAAAGTTCTTGGCGGCATGGATCTAATATCCAAGTTGATGA 1752
Qy 601 AGTCAGAGGCTGGTATATTTTGAAGGCCACCAAGACTCCCTTTTAGAGCATCACCTGTA 660
Db 1753 AGTCAGAGGCTGGTATATTTTGAAGGCCACCAAGACTCCCTTTTAGAGCATCACCTGTA 1812
Qy 661 CGTAGTCAGTTACGTAATCTTGGAGAGGTGACAAAGGCTGACTGACCGTGGCTACTCACA 720
Db 1813 CGTAGTCAGTTACGTAATCTTGGAGAGGTGACAAAGGCTGACTGACCGTGGCTACTCACA 1872
Qy 721 TCTTGTGTCATCAGTCAGCAGCTGTGACTTCTTTAAGTAAGTATAGTAACCAAGAA 780
Db 1873 TCTTGTGTCATCAGTCAGCAGCTGTGACTTCTTTAAGTAAGTATAGTAACCAAGAA 1932
Qy 781 TCCACACTGTGTCTCCCTTTCAAGCTATCAAGTCTTGAAGATGACCCCACTTTGCAAAAC 840
Db 1933 TCCACACTGTGTCTCCCTTTCAAGCTATCAAGTCTTGAAGATGACCCCACTTTGCAAAAC 1992
Qy 841 ABAGGAAATTTGGGCCACCATTTTGGATTTCAGAGGCTCTCTTCCGACTATACCTCC 900
Db 1993 ABAGGAAATTTGGGCCACCATTTTGGATTTCAGAGGCTCTCTTCCGACTATACCTCC 2052
Qy 901 AGAAATTTCTCTTTGAAGTACTCTGGAATTTACATGTATGGAATGCTCTCAAGGC 960
Db 2053 AGAAATTTCTCTTTGAAGTACTCTGGAATTTACATGTATGGAATGCTCTCAAGGC 2112
Qy 961 TCATGATCTACAGCTGGAAAGAAATATCCTACTGTGCTGTTTATATATGAGGCTCTCA 1020
Db 2113 TCATGATCTACAGCTGGAAAGAAATATCCTACTGTGCTGTTTATATATGAGGCTCTCA 2172
Qy 1021 ----- 1020
Db 2173 GGTGCAAGTGGTAATAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCCTAGC 2232
Qy 1021 ----- 1020
Db 2233 CTCCTAGGTTATGTGGTTGTAGTAGATAGACAACAGGGGATCCTGTCAACGAGGCTTAA 2292
Qy 1021 ----- 1020
Db 2293 ATTTGAAGGCGCTTTAAATATAAAATGGGTCAAAATAGAAATTCAGCATCAGGTGAAG 2352
Qy 1021 ----- 1020
Db 2353 ACTCCAATATCTAGCTTCTCGATATGATTTCAITTAGCTAGATCGTGTGGGCATCCACG 2412
Qy 1021 ----- 1020
Db 2413 CTGGTCTATGAGGATACCTCTCCCTGATGGCATTAATGCAGAGGTTCAGATATCTTCAG 2472
Qy 1021 GGTTCGTTATGTGGGGCCAGTCACCTCTGTGATCTTCTATGATACAGGATACACGGA 1080
Db 2473 GGTTCGTTATGTGGGGCCAGTCACCTCTGTGATCTTCTATGATACAGGATACACGGA 2532

Qy 1081 ACCTTATATGGGTCAACCTGTGACCAAGATGACAGGGCTATTACTTAGGATCTGTGGCAT 1140
Db 2533 ACCTTATATGGGTCAACCTGTGACCAAGATGACAGGGCTATTACTTAGGATCTGTGGCAT 2592
Qy 1141 GCAAGCAGAAAAGTTCCCTCTGAAACCAAAATCGTTTACTGTCTTTACATGATGTTTCC 1200
Db 2593 GCAAGCAGAAAAGTTCCCTCTGAAACCAAAATCGTTTACTGTCTTTACATGATGTTTCC 2652
Qy 1201 TGAGAAATGTCCTATTTTGCACATACCAAGTATATTTACTGTAGTTTTTTAGTGAGGCTG 1260
Db 2653 TGAGAAATGTCCTATTTTGCACATACCAAGTATATTTACTGTAGTTTTTTAGTGAGGCTG 2712
Qy 1261 GCCATATGATTTACAGATCTATCCTCAGGAGAGACACAGCATAAAGATTCCTGAATCG 1320
Db 2713 GCCATATGATTTACAGATCTATCCTCAGGAGAGACACAGCATAAAGATTCCTGAATCG 2772
Qy 1321 AGAACATTTAGAACTGCATCTTTTGCACATACCTTCAAGAAAAACCTTGGATCACGTTAT 1380
Db 2773 AGAACATTTAGAACTGCATCTTTTGCACATACCTTCAAGAAAAACCTTGGATCACGTTAT 2832
Qy 1381 TGCTCTAAAAGTGATATAATTTTGCACCTGTGTAGAACTCTCTGGTATACACTGGCTATTT 1440
Db 2833 TGCTCTAAAAGTGATATAATTTTGCACCTGTGTAGAACTCTCTGGTATACACTGGCTATTT 2892
Qy 1441 AACCAATGAGGAGGTTTAAATCAACAGAAAAACACAGAAATTCATCATCACTTTTGATACC 1500
Db 2893 AACCAATGAGGAGGTTTAAATCAACAGAAAAACACAGAAATTCATCATCACTTTTGATACC 2952
Qy 1501 TGCCATGTAACATCTACTCTCTGAAATAAATGTGTGTCATGCAAGGCTCTACGTTTGT 1560
Db 2953 TGCCATGTAACATCTACTCTCTGAAATAAATGTGTGTCATGCAAGGCTCTACGTTTGT 3012
Qy 1561 GGTAGTAATCTAATACCTTAACCCACATGCTCAAAATCAAAATGATATATCTCTGAGA 1620
Db 3013 GGTAGTAATCTAATACCTTAACCCACATGCTCAAAATCAAAATGATATATCTCTGAGA 3072
Qy 1621 GACCGAGCATACCATAGAAATTTACTAAAAAAA 1654
Db 3073 GACCGAGCATACCATAGAAATTTACTAAAAAAA 3106

RESULT 13
US-09-976-674-22
; Sequence 22, Application US/09976674
; Patent No. US20020115843A1
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 22
; LENGTH: 4685
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-976-674-22

Query Match 80.5%; Score 1343.6; DB 3; Length 4685;
Best Local Similarity 87.0%; Pred. No. 0;
Matches 1589; Conservative 0; Mismatches 79; Indels 158; Gaps 4;
Qy 1 AACAGGTACAGCAAAATCTTAAAGTCACCTTTAAAGATGTCAGAAATATATGATGCTGA 60
Db 1164 AACAGGTACAGCAAAATCTTAAAGTCACCTTTAAAGATGTCAGAAATATATGATGCTGA 1223
Qy 61 AGGAAGATCATAGATGTCATAGATAGGAACCTTAATCAACCTTTTGAGATTTCTATTGA 120

Best Local Similarity 87.0%; Pred. No. 0; Matches 1589; Conservative 0; Mismatches 79; Indels 158; Gaps 4;									
Qy	1	AACAGGTACAGCAAAATCCATAAGTCACTTTTAAAGATGTCAGAAATATGATTGATGCTGA	60						
Db	1164	AACAGGTACAGCAAAATCCATAAGTCACTTTTAAAGATGTCAGAAATATGATTGATGCTGA	1223						
Qy	61	AGGAAGGATCATAGATGTCATAGATAAGGAATCAATTCACCTTTTGGAGATTCCTATTGGA	120						
Db	1224	AGGAAGGATCATAGATGTCATAGATAAGGAATCAATTCACCTTTTGGAGATTCCTATTGGA	1283						
Qy	121	AGGAGTTGAATATATTGCCAGAGCTGGATGGAATCCTCAGGGGAAATATCTGGTCCAT	180						
Db	1284	AGGAGTTGAATATATTGCCAGAGCTGGATGGAATCCTCAGGGGAAATATCTGGTCCAT	1343						
Qy	181	CCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAATATTAT	240						
Db	1344	CCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAATATTAT	1403						
Qy	241	CCAGTAGAAGATGATGTTATGGAAAGGAGAGACTCAATTGAGTCAGTGCCTGATCTGT	300						
Db	1404	CCAGTAGAAGATGATGTTATGGAAAGGAGAGACTCAATTGAGTCAGTGCCTGATCTGT	1463						
Qy	301	GAGGCCATAATATCTATGAAAGAAACAAACAGACATCTGGATAAATATCCATGACATCTT	360						
Db	1464	GAGGCCATAATATCTATGAAAGAAACAAACAGACATCTGGATAAATATCCATGACATCTT	1523						
Qy	361	TCATGTTTTTCCCAAAGTCAGAAAGAGGAAATTTGAGTTATTTTTGCTCTGAAATGCAA	420						
Db	1524	TCATGTTTTTCCCAAAGTCAGAAAGAGGAAATTTGAGTTATTTTTGCTCTGAAATGCAA	1583						
Qy	421	AACAGGTTTCCGTCATTTATACAAATTTACATCTATTATTTAAAGGAAGCAATATATAACG	480						
Db	1584	AACAGGTTTCCGTCATTTATACAAATTTACATCTATTATTTAAAGGAAGCAATATATAACG	1643						
Qy	481	ATCCAGTGTGGGCTGCCTGCTCCAAAGTATTTCAAGTGTCCTATCAAGAGGAGATAGC	540						
Db	1644	ATCCAGTGTGGGCTGCCTGCTCCAAAGTATTTCAAGTGTCCTATCAAGAGGAGATAGC	1703						
Qy	541	AATTACAGTGTGATGGGAATTTTGAAGGCAACCAAGACTCCCTTTAGAGCATCACTGTA	600						
Db	1704	AATTACAGTGTGATGGGAATTTTGAAGGCAACCAAGACTCCCTTTAGAGCATCACTGTA	1763						
Qy	601	AGTCAGAGGCTGTATATTTTGAAGGCAACCAAGACTCCCTTTAGAGCATCACTGTA	660						
Db	1764	AGTCAGAGGCTGTATATTTTGAAGGCAACCAAGACTCCCTTTAGAGCATCACTGTA	1823						
Qy	661	CGTAGTCAGTTACGTAATCTCTGGAGAGGTGACAGGCTGACCTGCTTTCAGAGCATCACTGTA	720						
Db	1824	CGTAGTCAGTTACGTAATCTCTGGAGAGGTGACAGGCTGACCTGCTTTCAGAGCATCACTGTA	1883						
Qy	721	TTCTTGCTGATCAGTCAGCACTGTGACTTTTATAAGTAAGTATAGTAACCAAGAGAA	780						
Db	1884	TTCTTGCTGATCAGTCAGCACTGTGACTTTTATAAGTAAGTATAGTAACCAAGAGAA	1943						
Qy	781	TCCACACTGTGTGCTCCCTTTTCAAGCTATCAAGTCTCTGAAGATGACCCCACTTGCAAAAC	840						
Db	1944	TCCACACTGTGTGCTCCCTTTTCAAGCTATCAAGTCTCTGAAGATGACCCCACTTGCAAAAC	2003						
Qy	841	AAAGAAATTTGGGCCACCAATTTTGGATTTCAG---CAGTCTCTTCTCTGACTATATCTC	896						
Db	2004	AAAGAAATTTGGGCCACCAATTTTGGATTTCAGTCTCTGAGTGTGATGATATCGG	2063						
Qy	897	CTCCA-----GAAATTTCTCTTTTGAAGTACTACTGGATTACATGATGGATGTC	950						
Db	2064	TTTAAAGGAGTCAAGTATTTCCGGCTGATACCTAGCTCTCTAGTTTATGTTGTTGA	2123						
Qy	951	TCTCAAGCCTCATGATCTACAG-----	973						
Db	2124	GTGATAGACACAGGGATCTGTCCACGAGGCTTAAATTTGAAGGGCCCTTTAAATAT	2183						
Qy	974	-----	973						

Db	2184	AAATGGGTCAATAGAAATTCAGCATCAGGTGGAAGNACTCCAATATCTACGTTCTCGA	2243
Qy	974	-----CCTGGAAAGAAATATCCTACTGTGCTGTCTATATATGTTGGTCTCTC	1019
Db	2244	TATGATTTCAATGACCTAGATCGTGTGGGCATCCAGGCTGTCTCTATGAGGATACCTC	2303
Qy	1020	-----AGTTTGTCTATTGCTTGGGGCCCA	1042
Db	2304	TCCCTGATGCATTAATGACAGGCTCAGATATCTTCAGGGTTGCTATTGCTGGGGCCCA	2363
Qy	1043	GTCACTCTGFGATCTTCTATGATACAGGATACACGGAACGTTATATGGGTCACCTGAC	1102
Db	2364	GTCACTCTGFGATCTTCTATGATACAGGATACACGGAACGTTATATGGGTCACCTGAC	2423
Qy	1103	CAGAATGAACAGGGCTATTACTTAGGATCTGTGGCATGCAAGCAGAGAAAGTTCCCTCT	1162
Db	2424	CAGAATGAACAGGGCTATTACTTAGGATCTGTGGCATGCAAGCAGAGAAAGTTCCCTCT	2483
Qy	1163	GAACCAAAATCGTTTACTGCTCTTACATGTTTCTGTGATGAGAAATGTCATTTTGACAT	1222
Db	2484	GAACCAAAATCGTTTACTGCTCTTACATGTTTCTGTGATGAGAAATGTCATTTTGACAT	2543
Qy	1223	ACCAGTATATTAATGAGTTTCTGAGGCTGGAAAGCCATATGATTTACAGATCTAT	1282
Db	2544	ACCAGTATATTAATGAGTTTCTGAGGCTGGAAAGCCATATGATTTACAGATCTAT	2603
Qy	1283	CCTCAGGAGACACAGCATAGAGTTCTGATCGGATGAGAAATGTCATTTGACATCTT	1342
Db	2604	CCTCAGGAGACACAGCATAGAGTTCTGATCGGATGAGAAATGTCATTTGACATCTT	2663
Qy	1343	TTGCACTACTCTCAAGAAACCTTGATCAGTATGCTGCTCTTAAAGTGATATAATTT	1402
Db	2664	TTGCACTACTCTCAAGAAACCTTGATCAGTATGCTGCTCTTAAAGTGATATAATTT	2723
Qy	1403	TGACCTGTGTAGAACTCTCTGTATACACTGGCTATTTAAACAAATGAGGAGTTAAATC	1462
Db	2724	TGACCTGTGTAGAACTCTCTGTATACACTGGCTATTTAAACAAATGAGGAGTTAAATC	2783
Qy	1463	AACAGAAACACAGAAATGATCATCATATTTGATACCTGCCATGTAACCTACTCTCTG	1522
Db	2784	AACAGAAACACAGAAATGATCATCATATTTGATACCTGCCATGTAACCTACTCTCTG	2843
Qy	1523	AAATAAATGTGGTCCCATGAGGGGTCTACGGTTTGTGGTAGTAATCTAATACCTTAAC	1582
Db	2844	AAATAAATGTGGTCCCATGAGGGGTCTACGGTTTGTGGTAGTAATCTAATACCTTAAC	2903
Qy	1583	CCCATGCTCAAAATCAAAATGATACATATTCCTTGAGAGCCAGCAATACCAATAAGAT	1642
Db	2904	CCCATGCTCAAAATCAAAATGATACATATTCCTTGAGAGCCAGCAATACCAATAAGAT	2963
Qy	1643	TACTAAAAAATAAAAAAAAAAAAAA	1668
Db	2964	TACTAAAAAATAAAAAAAAAAAAAA	2989

RESULT 15
US-10-275-505-16
; Sequence 15, Application US/10275505
; Publication No. US20040081961A1
; GENERAL INFORMATION
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: BELGEBANE, Angelo M.; LAL, Preeti G.
; APPLICANT: WALIA, April J.A.; PATTERSON, Chandra
; APPLICANT: WALIA, Narinder K.; KEARNEY, Liam
; APPLICANT: TRIBOULEY, Catherine M.; KHAN, Farrah A.
; APPLICANT: YAO, Monique G.; BAUGHN, Mariah R.
; APPLICANT: AZIMZAI, Yalda; ELLIOTT, Vicki S.
; APPLICANT: NGUYEN, Dannel B.; GANDHI, Ameena R.
; APPLICANT: YANG, Junning; HERNANDEZ, Roberto
; APPLICANT: POLICKY, Jennifer L.; LU, Dyung Aina M.
; APPLICANT: REDDY, Roopa M.; YUE, Henry
; APPLICANT: TANG, Y. Tom
; TITLE OF INVENTION: PROTEASES

FILE REFERENCE: PI-0085 USN
CURRENT APPLICATION NUMBER: US/10/275,505
CURRENT FILING DATE: 2002-11-04
PRIOR APPLICATION NUMBER: PCT/US01/14651
PRIOR FILING DATE: 2001-05-04
PRIOR APPLICATION NUMBER: 60/209,402
PRIOR FILING DATE: 2000-06-01
PRIOR APPLICATION NUMBER: 60/207,477
PRIOR FILING DATE: 2000-05-25
PRIOR APPLICATION NUMBER: 60/205,803
PRIOR FILING DATE: 2000-05-17
PRIOR APPLICATION NUMBER: 60/203,566
PRIOR FILING DATE: 2000-05-11
PRIOR APPLICATION NUMBER: 60/202,082
PRIOR FILING DATE: 2000-05-04
NUMBER OF SEQ ID NOS: 28
SOFTWARE: PERL Program
SEQ ID NO 16
LENGTH: 2510
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No: 376067CB1
US-10-275-505-16

Query Match 74.9%; Score 1250.6; DB 7; Length 2510;
Best Local Similarity 87.6%; Pred. No. 0;
Matches 1482; Conservative 0; Mismatches 44; Indels 165; Gaps 4;

Qy	1	AACAGGTACAGCAAAATCTTAAAGTCACTTTTAAAGATGTCAGAAATAATGATGCTGA	60
Db	831	AACAGGTACAGCAAAATCTTAAAGTCACTTTTAAAGATGTCAGAAATAATGATGCTGA	890
Qy	61	AGGAAGGATCATAGATGTCATAGATAAGGAACATAATCAACCTTTTGAGATTCATTGA	120
Db	891	AGGAAGGATCATAGATGTCATAGATAAGGAACATAATCAACCTTTTGAGATTCATTGA	950
Qy	121	AGGAGTGAATATATTTGCGAGCTGGATGGATCTCTGAGGGAATAATGCTGGTCCAT	180
Db	951	AGGAGTGAATATATTTGCGAGCTGGATGGATCTCTGAGGGAATAATGCTGGTCCAT	1010
Qy	181	CTTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAATTTAT	240
Db	1011	CTTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAATTTAT	1070
Qy	241	CCAGTAGAAGATGATGTTATGGAAGGCAGAGACTCAATGAGTCAGTGCCTGATTCGT	300
Db	1071	CCAGTAGAAGATGATGTTATGGAAGGCAGAGACTCAATGAGTCAGTGCCTGATTCGT	1130
Qy	301	GAGCCCACTAATTTATCTATGAAGAAACAACAGACATCTGGATAAATATCCATGACATCT	360
Db	1131	GAGCCCACTAATTTATCTATGAAGAAACAACAGACATCTGGATAAATATCCATGACATCT	1190
Qy	361	TCATGTTTTTCCCAAGTCAGAGAGGAATTCAGTTTATTTTGGCTCTGATGCAA	420
Db	1191	TCATGTTTTTCCCAAGTCAGAGAGGAATTCAGTTTATTTTGGCTCTGATGCAA	1250
Qy	421	AACAGGTTCCGTCATTATACAAATTCATCTATTTTAAAGGAACAAATAAAG	480
Db	1251	AACAGGTTCCGTCATTATACAAATTCATCTATTTTAAAGGAACAAATAAAG	1310
Qy	481	ATCCAGTGGGGCTGGCTCCAGTCAATTTCAAGTGTCTATCAAGAGGAGATAGC	540
Db	1311	ATCCAGTGGGGCTGGCTCCAGTCAATTTCAAGTGTCTATCAAGAGGAGATAGC	1369
Qy	541	AATTTACCAGTGGTGAATGGGAAGTTCTTGGCCGGATGGATCTAATATCCAAAGTTGATGA	600
Db	1370	TAGGAAC-----TCCATCTGTATGTGTGTGACATATATGTTGAGATCCAAAGTTGATGA	1424
Qy	601	AGTCAGAGGCTGGTATATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACCTGTA	660
Db	1425	AGTCAGAGGCTGGTATATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACCTGTA	1484

Qy	661	CGTAGTCAGTTACGTAAATCCCTGGAGAGGTGACAAAGGCTGACTGACCGTGGCTACTCACA	720
Db	1485	CGTAGTCAGTTACGTAAATCCCTGGAGAGGTGACAAAGGCTGACTGACCGTGGCTACTCACA	1544
Qy	721	TTCTTGTGTCATCAGTCAGCACTGTGACTTCTTTTAAAGTAAGTATAGTATACCAAGAGAA	780
Db	1545	TTCTTGTGTCATCAGTCAGCACTGTGACTTCTTTTAAAGTAAGTATAGTATACCAAGAGAA	1604
Qy	781	TCACACTGTGTCTCCCTTTTCAAGCTATCAAGTCTCAAGATGAGATGCCAATCTTGCAGAAC	840
Db	1605	TCACACTGTGTCTCCCTTTTCAAGCTATCAAGTCTCAAGATGAGATGCCAATCTTGCAGAAC	1664
Qy	841	AAAGGAATTTTGGCCACCATTCTTGGATTCAGACAGGTCTCTTCTGCTGACTATATCTCTCC	900
Db	1665	AAAGGAATTTTGGCCACCATTCTTGGATTCAGACAGGTCTCTCTTCTGCTGACTATATCTCTCC	1724
Qy	901	AGAAATTTTCTCTTTTGAAGTACTACTGATTTTACATTTGATGGATGCTCTACAAGCC	960
Db	1725	AGAAATTTTCTCTTTTGAAGTACTACTGATTTTACATTTGATGGATGCTCTACAAGCC	1784
Qy	961	TCATGATCTACAGCCTGGAAAGAAATATCTTACTGTCTGCTGTTTCATATATGGTGGCTCTCA	1020
Db	1785	TCATGATCTACAGCCTGGAAAGAAATATCTTACTGTCTGCTGTTTCATATATGGTGGCTCTCA	1844
Qy	1021	-----	1020
Db	1845	GGTGCAGTTGGTGAATAATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCCTAGC	1904
Qy	1021	-----	1020
Db	1905	CTCTCTAGTTTGTGTTGTAGTGATGACAAACAGGGGATCCTGTCAACGAGGGCTTAA	1964
Qy	1021	-----GGTTCCTATGCTGGGGCCCGAGTCACCTCTGTG	1053
Db	1965	ATTTGAAGCGCCTTTAAATATATAAATGTTGCTATGCTGGGGCCCGAGTCACCTCTGTG	2024
Qy	1054	GATCTTCTATGATACAGGATACACGGAAGTTTATATGGGTCAACCTGACCAAGATGAACA	1113
Db	2025	GATCTTCTATGATACAGGATACACGGAAGTTTATATGGGTCAACCTGACCAAGATGAACA	2084
Qy	1114	GGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTGAAACCAATFCG	1173
Db	2085	GGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTGAAACCAATFCG	2144
Qy	1174	TTTACTGCTCTTACATGTTTCTGTGATGAAATGTCATTGTCACATACCAAGTATAT	1233
Db	2145	TTTACTGCTCTTACATGTTTCTGTGATGAAATGTCATTGTCACATACCAAGTATAT	2204
Qy	1234	ACTGAGTTTTTATGAGGGCTGGAAAGCCATATGATTTACAGATCTATCTCAGGAGAG	1293
Db	2205	ACTGAGTTTTTATGAGGGCTGGAAAGCCATATGATTTA-----CAGGAGAG	2252
Qy	1294	ACAGAGCATAGAGTTCCCTGAATCGGAGAAACATTATGAACTGCACTTTTGGCACTACT	1353
Db	2253	ACAGAGCATAGAGTTCCCTGAATCGGAGAAACATTATGAACTGCACTTTTGGCACTACT	2312
Qy	1354	TCAAGAAAACCTTGGATCACGTTATGCTGCTCTTAAAGTGAATATAATTTTGACCTGTGA	1413
Db	2313	TCAAGAAAACCTTGGATCACGTTATGCTGCTCTTAAAGTGAATATAATTTTGACCTGTGA	2372
Qy	1414	GAACTCTCTGGTATACATGCTGCTTAAACAAATAGGAGGTTTAAATCAACAGAAAAACA	1473
Db	2373	GAACTCTCTGGTATACATGCTGCTTAAACAAATAGGAGGTTTAAATCAACAGAAAAACA	2432
Qy	1474	CAGAAATGATCATCATATTTTGATACCTGCAATGTAACTACTCTCTGAAATAAATGT	1533
Db	2433	CAGAAATGATCATCATATTTTGATACCTGCAATGTAACTACTCTCTGAAATAAATGT	2492
Qy	1534	GGTGCCATGCA	1544
Db	2493	GGTGCCATGAA	2503

Search completed: May 4, 2006, 11:21:52
Job time : 1496.73 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 4, 2006, 03:49:39 ; Search time 2279.71 Seconds
(without alignments)
2981.544 Million cell updates/sec

Title: US-10-825-632-6
Perfect score: 1669
Sequence: 1 aaacgtagcagcaaatccta.....aaaaaaaaaaaaaaaaaa 1669

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 9306428 seqs, 2036268586 residues

Total number of hits satisfying chosen parameters: 18612856

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA_New:
1: /SIDSS/ptodata/2/pubpna/US08_NEW_PUB.seq1:
2: /SIDSS/ptodata/2/pubpna/US06_NEW_PUB.seq:
3: /SIDSS/ptodata/2/pubpna/US07_NEW_PUB.seq:
4: /SIDSS/ptodata/2/pubpna/US08_NEW_PUB.seq:
5: /SIDSS/ptodata/2/pubpna/PCT_NEW_PUB.seq:
6: /SIDSS/ptodata/2/pubpna/US03_NEW_PUB.seq:
7: /SIDSS/ptodata/2/pubpna/US09_NEW_PUB.seq:
8: /SIDSS/ptodata/2/pubpna/US10_NEW_PUB.seq:
9: /SIDSS/ptodata/2/pubpna/US10_NEW_PUB.seq:
10: /SIDSS/ptodata/2/pubpna/US10_NEW_PUB.seq:
11: /SIDSS/ptodata/2/pubpna/US10_NEW_PUB.seq:
12: /SIDSS/ptodata/2/pubpna/US10_NEW_PUB.seq:
13: /SIDSS/ptodata/2/pubpna/US10_NEW_PUB.seq:
14: /SIDSS/ptodata/2/pubpna/US11_NEW_PUB.seq:
15: /SIDSS/ptodata/2/pubpna/US11_NEW_PUB.seq:
16: /SIDSS/ptodata/2/pubpna/US11_NEW_PUB.seq:
17: /SIDSS/ptodata/2/pubpna/US11_NEW_PUB.seq:
18: /SIDSS/ptodata/2/pubpna/US11_NEW_PUB.seq:
19: /SIDSS/ptodata/2/pubpna/US60_NEW_PUB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1353.4	81.1	3143	17	US-11-151-601-19
2	1087.4	65.2	2649	17	US-11-151-601-21
C 3	118	7.1	1346	7	US-09-925-065A-669313
C 4	113.6	6.8	609	7	US-09-925-065A-743558
C 5	113.6	6.8	609	7	US-09-925-065A-743559
C 6	73.4	4.4	612	7	US-09-925-065A-818446
7	51.6	3.1	2238	18	US-11-079-463-1186
8	50.6	3.0	378	11	US-10-932-182A-81332
9	50.6	3.0	378	11	US-10-932-182A-81332
10	49.6	3.0	2778	11	US-10-932-182A-5649
11	49.6	3.0	2778	11	US-10-932-182A-5649
12	45.6	2.7	332	18	US-11-208-288-1
13	45.6	2.7	3407	11	US-10-501-035-34
14	44.8	2.7	1624	10	US-10-131-826A-181

15	44.8	2.7	1624	11	US-10-973-115B-181	Sequence 181, App
16	44.8	2.7	1624	13	US-10-137-873A-181	Sequence 181, App
17	44.8	2.7	1624	13	US-10-152-370-181	Sequence 181, App
18	44.8	2.7	1624	18	US-11-290-153-181	Sequence 181, App
19	44	2.6	2217	18	US-11-208-288-3	Sequence 3, Appli
20	44	2.6	2301	10	US-10-522-789-1	Sequence 1, Appli
21	42	2.5	2283	18	US-11-208-288-5	Sequence 5, Appli
22	41	2.5	588	7	US-09-925-065A-486279	Sequence 486279, A
23	41	2.5	622	7	US-09-925-065A-49906	Sequence 49906, A
24	41	2.5	622	11	US-10-301-480-151144	Sequence 151144, A
25	41	2.5	622	12	US-10-301-480-764553	Sequence 764553, A
26	40.4	2.4	600	17	US-11-136-527-6535	Sequence 6535, Ap
27	40.4	2.4	1825	17	US-11-136-527-2439	Sequence 2439, Ap
28	40	2.4	660	13	US-10-194-487-467	Sequence 467, App
29	40	2.4	660	13	US-10-195-883-467	Sequence 467, App
30	40	2.4	660	13	US-10-195-888-467	Sequence 467, App
31	40	2.4	660	13	US-10-195-889-467	Sequence 15, Appl
32	40	2.4	660	13	US-10-226-486-15	Sequence 179, App
33	40	2.4	666	13	US-10-986-405-179	Sequence 143, App
34	40	2.4	670	13	US-10-986-405-143	Sequence 3494, Ap
35	39.8	2.4	600	10	US-10-750-185-3494	Sequence 3494, Ap
36	39.8	2.4	600	10	US-10-750-623-3494	Sequence 2130, Ap
37	39.8	2.4	4852	17	US-11-136-527-2130	Sequence 47, Appl
C 38	39.8	2.4	5152	10	US-10-240-708-47	Sequence 35, Appl
C 39	39.8	2.4	170995	17	US-11-121-086-35	Sequence 6859, Ap
40	39.6	2.4	1400	17	US-11-136-527-6859	Sequence 2763, Ap
41	39.6	2.4	3708	17	US-11-136-527-2763	Sequence 394749, A
42	39.4	2.4	555	7	US-09-925-065A-394749	Sequence 579, App
43	39.4	2.4	1516	9	US-10-505-928-579	Sequence 61, Appl
44	39.4	2.4	1944	13	US-10-782-413-61	Sequence 38461, A
C 45	39	2.3	1819	10	US-10-750-185-38461	

ALIGNMENTS

RESULT 1
US-11-151-601-19
; Sequence 19, Application US/1151601
; Publication No. US20060003413A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Olandt, Peter J.
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Curtis, Rory A. J.
; APPLICANT: Williamson, Mark
; APPLICANT: Weich, Nadine
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASE, PHOSPHATASE
; FILE REFERENCE: AND PROTEASE FAMILY MEMBERS AND USES THEREOF
; CURRENT APPLICATION NUMBER: US/11/151,601
; CURRENT FILING DATE: 2005-06-13
; PRIOR APPLICATION NUMBER: US 10/170,789
; PRIOR FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: US 09/797,039
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: PCT/US01/06525
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/186,061
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 09/882,166
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: PCT/US01/19269
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/212,078
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 09/934,406
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: PCT/US01/26052
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/226,740
; PRIOR FILING DATE: 2000-08-21

```
/ Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 3143
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (229)....(2874)
US-11-151-601-19

Query Match      81.1%; Score 1353.4; DB 17; Length 3143;
Best Local Similarity 84.7%; Pred. No. 9.8e-289;
Matches 1664; Conservative 0; Mismatches 1; Indels 300; Gaps 1;

Qy 1 AACAGGTACAGCAAAATCCCTAAAGTCACTTTTAAAGATGTCAGAAATAATGATGCTGA 60
    |||
Db 1179 AACAGGTACAGCAAAATCCCTAAAGTCACTTTTAAAGATGTCAGAAATAATGATGCTGA 1238
    |||

Qy 61 AGGAAGATCATAGATGTCATAGATAAGGAACCTAATCAACCTTTTGAGATTCCTATTGA 120
    |||
Db 1239 AGGAAGATCATAGATGTCATAGATAAGGAACCTAATCAACCTTTTGAGATTCCTATTGA 1298
    |||

Qy 121 AGGAGTTGAATATATTGCCAGAGCTGGATGGAATCCTGAGGGAATAATGCTGGTCCAT 180
    |||
Db 1299 AGGAGTTGAATATATTGCCAGAGCTGGATGGAATCCTGAGGGAATAATGCTGGTCCAT 1358
    |||

Qy 181 CCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAATTTTAT 240
    |||
Db 1359 CCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAATTTAT 1418
    |||

Qy 241 CCCAGTAGAAGATGATGTTATGGAAGGCAGAGACTCAATGAGTCAGTGCCTGCTCTGT 300
    |||
Db 1419 CCCAGTAGAAGATGATGTTATGGAAGGCAGAGACTCAATGAGTCAGTGCCTGCTCTGT 1478
    |||

Qy 301 GAGCCACATTAATCTATGAGAAACAACAGACATCTGGAATAATTCATGACATCTT 360
    |||
Db 1479 GAGCCACATTAATCTATGAGAAACAACAGACATCTGGAATAATTCATGACATCTT 1538
    |||

Qy 361 TCATGTTTTCCCAAGTCAGAGAGAAATGAGTTTATTTTGGCTCTGAATGCAA 420
    |||
Db 1539 TCATGTTTTCCCAAGTCAGAGAGAAATGAGTTTATTTTGGCTCTGAATGCAA 1598
    |||

Qy 421 AACAGGTTCCGTCAATTTATACAAATTAATCATCTATTTTAAAGAAAGCAAAATATAAAG 480
    |||
Db 1599 AACAGGTTCCGTCAATTTATACAAATTAATCATCTATTTTAAAGAAAGCAAAATATAAAG 1658
    |||

Qy 481 ATCCAGTGTGGGCTGCTGCTCCAAAGTGATTTCAAAGTGCTCTATCAAAGAGAGATAGC 540
    |||
Db 1659 ATCCAGTGTGGGCTGCTGCTCCAAAGTGATTTCAAAGTGCTCTATCAAAGAGAGATAGC 1718
    |||

Qy 541 AATTACAGTGTGAATGGGAAGTTCTTGGCCGCGCATGGATCTAATATCCAAGTTGATGA 600
    |||
Db 1719 AATTACAGTGTGAATGGGAAGTTCTTGGCCGCGCATGGATCTAATATCCAAGTTGATGA 1778
    |||

Qy 601 AGTCAGAGGCTGGTATATTTTGAAGGCACCAAAAGACTCCCTTTAGAGCATCACCTGTA 660
    |||
Db 1779 AGTCAGAGGCTGGTATATTTTGAAGGCACCAAAAGACTCCCTTTAGAGCATCACCTGTA 1838
    |||

Qy 661 CGTAGTCAGTTACGTAATCTCTGGAGAGGTGACAAAGCTGACTGACCGTGGCTACTACA 720
    |||
Db 1839 CGTAGTCAGTTACGTAATCTCTGGAGAGGTGACAAAGCTGACTGACCGTGGCTACTACA 1898
    |||

Qy 721 TTCTTGTCATCAGTCAGCAGCTGACTTCTTTTAAAGTAAGTATAGTAACAGAGAA 780
    |||
Db 1899 TTCTTGTCATCAGTCAGCAGCTGACTTCTTTTAAAGTAAGTATAGTAACAGAGAA 1958
    |||

Qy 781 TCCACACTGTGTGCTCCCTTACAGCTATCAAGTCTGAGATGACCAACTGCAAAAC 840
    |||
Db 1959 TCCACACTGTGTGCTCCCTTACAGCTATCAAGTCTGAGATGACCAACTGCAAAAC 2018
    |||

Qy 841 AAAGGAAATTTGGGCCCAACATTTTGGATTGAGAGGTCCTCTCTCTGACTATACCTCTCC 900
    |||
```

```
Db 2019 AAAGGAAATTTGGGCCCAACATTTTGGATTGAGAGGTCCTCTCTCTGACTATACCTCTCC 2078
    |||
Qy 901 AGAAATTTCTCTTTTGAAGTACTACTGGAATTTACATTTGATGGGATGCTCTACAAGCC 960
    |||
Db 2079 AGAAATTTCTCTTTTGAAGTACTACTGGAATTTACATTTGATGGGATGCTCTACAAGCC 2138
    |||
Qy 961 TCATGATCTACAGCCTGGAAAGAAATATCCTACTGCTGTGTTTCATATATGATGTCCTCA 1020
    |||
Db 2139 TCATGATCTACAGCCTGGAAAGAAATATCCTACTGCTGTGTTTCATATATGATGTCCTCA 2198
    |||
Qy 1021 ----- 1020
    |||
Db 2199 GGTGAGTGTGGTAATAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTAGC 2258
    |||
Qy 1021 ----- 1020
    |||
Db 2259 CTCTCTAGTTATGTGTGTTAGTAGACACAACAGGGGATCCTGTCAACGAGGGCTTAA 2318
    |||
Qy 1021 ----- 1020
    |||
Db 2319 ATTTGAAGGCGCCTTTAAATATAAAATGGGTCAAATAGAAATTTGACGATCAGGTGGAAG 2378
    |||
Qy 1021 ----- 1020
    |||
Db 2379 ACTCCAATATCTAGCTTCTCGATATGATTTTCATTGACTTAGATCGTGTGGGCATCCACGG 2438
    |||
Qy 1021 ----- 1020
    |||
Db 2439 CTGGTCTTGGAGGATACCTCTCCCTGATGGCATTAATGCAGAGGTCAGATATCTTCAG 2498
    |||
Qy 1021 GGTGTGCTNTGTCGGGCGCCAGTCATCTGTGATCTCTTATGATACAGGATACACGGA 1080
    |||
Db 2499 GGTGTGCTNTGTCGGGCGCCAGTCATCTGTGATCTCTTATGATACAGGATACACGGA 2558
    |||
Qy 1081 ACCTTATATGGGTCAACCTTGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCCAT 1140
    |||
Db 2559 ACCTTATATGGGTCAACCTTGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCCAT 2618
    |||
Qy 1141 GCAAGCAGAAAAAGTTCCCTCTGAAACCAAAATCGTTTTACTGCTCTTACATGGTTTCTCGGA 1200
    |||
Db 2619 GCAAGCAGAAAAAGTTCCCTCTGAAACCAAAATCGTTTTACTGCTCTTACATGGTTTCTCGGA 2678
    |||
Qy 1201 TGAGAAATGTCCATTTTGGCACATACCAAGTATATTACTAGAGTTTTTTAGTGAGGGCTGAAA 1260
    |||
Db 2679 TGAGAAATGTCCATTTTGGCACATACCAAGTATATTACTAGAGTTTTTTAGTGAGGGCTGAAA 2738
    |||
Qy 1261 GCCATATGATTTACAGATCTATCTTCAGGAGAGACACAGCATATAAGAGTTCCTGAATCGGG 1320
    |||
Db 2739 GCCATATGATTTACAGATCTATCTTCAGGAGAGACACAGCATATAAGAGTTCCTGAATCGGG 2798
    |||
Qy 1321 AGAACATTTAGAACTGCATCTTTTGCATCTA CTTTCAAGAAAAACCTTGGATCAGCTATTGC 1380
    |||
Db 2799 AGAACATTTAGAACTGCATCTTTTGCATCTA CTTTCAAGAAAAACCTTGGATCAGCTATTGC 2858
    |||
Qy 1381 TGCTCTAAAAGTGATATAATTTTGGACCTGTGTAGAACTCTCTGCTATACACTGGCTATTT 1440
    |||
Db 2859 TGCTCTAAAAGTGATATAATTTTGGACCTGTGTAGAACTCTCTGCTATACACTGGCTATTT 2918
    |||
Qy 1441 AACCAATCAGGAGGTTTAAATCAACAGAAACACAGAAATTTGATCATCATATTTTGTATACC 1500
    |||
Db 2919 AACCAATCAGGAGGTTTAAATCAACAGAAACACAGAAATTTGATCATCATATTTTGTATACC 2978
    |||
Qy 1501 TGCCATGTAACATCTACTCTCTGAAATAATTTGCTGGTCCATGCAAGGGCTCTACGGTTTGT 1560
    |||
Db 2979 TGCCATGTAACATCTACTCTCTGAAATAATTTGCTGGTCCATGCAAGGGCTCTACGGTTTGT 3038
    |||
Qy 1561 GGTAGTAATCTTAATACCTTAACCCCAATGCTCAAAATCAAAATGATATATTTCTCTGAGA 1620
    |||
Db 3039 GGTAGTAATCTTAATACCTTAACCCCAATGCTCAAAATCAAAATGATATATTTCTCTGAGA 3098
    |||
Qy 1621 GACCCAGCAATACCAATAAGAAATTTACTAAAAAATAAAAAA 1665
    |||
```

Db 3099 GA^{CC}CAGCAATACCATAGAATTACTAAAAA^{AAAAAAAAAAAAAAAAAAAA} 3143

RESULT 2

US-11-151-601-21
; Sequence 21, Application US/11151601
; Publication No. US20060003413A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Olandt, Peter J.
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Curtis, Rory A. J.
; APPLICANT: Williamson, Mark
; APPLICANT: Welch, Nadine
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASE, PHOSPHATASE,
; FILE OF INVENTION: AND PROTEASE FAMILY MEMBERS AND USES THEREOF
; FILE REFERENCE: MPI00-054P1RCP1OMNIDVIM
; CURRENT APPLICATION NUMBER: US/11/151,601
; CURRENT FILING DATE: 2005-06-13
; PRIOR APPLICATION NUMBER: US 10/170,789
; PRIOR FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: US 09/797,039
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: PCT/US01/06525
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/186,061
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 09/882,166
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: PCT/US01/19269
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/212,078
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 09/934,406
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: PCT/US01/26052
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/226,740
; PRIOR FILING DATE: 2000-08-21
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 2649
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-151-601-21

Query Match 65.2%; Score 1087.4; DB 17; Length 2649;
Best Local Similarity 82.3%; Pred. No. 5e-230;
Matches 1398; Conservative 0; Mismatches 1; Indels 300; Gaps 1;

QY 1 AACAGGTACAGCAATCTTAAGTCACCTTTTAAGATGTGACAGAAATAGATGATGCTGA 60
Db 951 AACAGGTACAGCAATCTTAAGTCACCTTTTAAGATGTGACAGAAATAGATGATGCTGA 1010
QY 61 AGGAGGATCATAGATGTCTAGATAAGGAACCTTAATCAACCTTTTGAGATTCATTATGA 120
Db 1011 AGGAGGATCATAGATGTCTAGATAAGGAACCTTAATCAACCTTTTGAGATTCATTATGA 1070
QY 121 AGGAGTGAATATATTTGACAGCTGGATGGATCTCTGAGGAAATATGCTTGCTCCAT 180
Db 1071 AGGAGTGAATATATTTGACAGCTGGATGGATCTCTGAGGAAATATGCTTGCTCCAT 1130
QY 181 CCTACTAGATCGCTCCAGACTCGCCTACAGATAGTGTGATCTCACCTGAAATATTTAT 240
Db 1131 CCTACTAGATCGCTCCAGACTCGCCTGAGATAGTGTGATCTCACCTGAAATATTTAT 1190
QY 241 CCCAGTAGAAGATGATGTTATGAAAGGACAGAGACTCATTTGAGTCAGTGCCTGATTCGT 300
Db 1191 CCCAGTAGAAGATGATGTTATGAAAGGACAGAGACTCATTTGAGTCAGTGCCTGATTCGT 1250

QY 301 GACGCCACTAATTATCTATGAAGAAACACAGACATCTGGATAAATATCCATGACATCTT 360
Db 1251 GACGCCACTAATTATCTATGAAGAAACACAGACATCTGGATAAATATCCATGACATCTT 1310
QY 361 TCATGTTTTTCCCAAGTCAAGAGGAAATAGATTTATTTTAAAGGAAGCAATATAAAGC 480
Db 1311 TCATGTTTTTCCCAAGTCAAGAGGAAATAGATTTATTTTAAAGGAAGCAATATAAAGC 1370
QY 421 AACAGTTTCGTCATTTATACAAATATACATCTATTTTAAAGGAAGCAATATAAAGC 480
Db 1371 AACAGTTTCGTCATTTATACAAATATACATCTATTTTAAAGGAAGCAATATAAAGC 1430
QY 481 ATCCAGTGTGGTGGCTGCTCTCAAGTGTATTTCAAGTGTCTCTCAAGAGGAGATAGC 540
Db 1431 ATCCAGTGTGGTGGCTGCTCTCAAGTGTATTTCAAGTGTCTCTCAAGAGGAGATAGC 1490
QY 541 AATTACCAAGTGTGAATGGGAAGTTCTTGCCCGGCATGGATCTAATATCCAAAGTTGATGA 600
Db 1491 AATTACCAAGTGTGAATGGGAAGTTCTTGCCCGGCATGGATCTAATATCCAAAGTTGATGA 1550
QY 601 AGTCAGAGGCTGGTATATTTTGAAGGACCAAGACTCCCTTTAGAGCATCACCTGTA 660
Db 1551 AGTCAGAGGCTGGTATATTTTGAAGGACCAAGACTCCCTTTAGAGCATCACCTGTA 1610
QY 661 CGTAGTCAGTTACGTAATCTCGAGAGGTGACAGGCTGACGTGACCGTGGCTACTCACA 720
Db 1611 CGTAGTCAGTTACGTAATCTCGAGAGGTGACAGGCTGACGTGACCGTGGCTACTCACA 1670
QY 721 TTCCTGTCGATCAGTCAGCAGCTGTGACTTCTTTTATAAGTAAAGTATAGTAAACAGAA 780
Db 1671 TTCCTGTCGATCAGTCAGCAGCTGTGACTTCTTTTATAAGTAAAGTATAGTAAACAGAA 1730
QY 781 TCCACACTGTGTGCTCCCTTTTACAAGCTATCAAGTGTCTGAAGATGACCCAACTTGCAAAC 840
Db 1731 TCCACACTGTGTGCTCCCTTTTACAAGCTATCAAGTGTCTGAAGATGACCCAACTTGCAAAC 1790
QY 841 AAAGGAATTTTGGGCCACCACTTTTGGATTTCAGCAGGTCTCTTCTCTGACTATACCTCTCC 900
Db 1791 AAAGGAATTTTGGGCCACCACTTTTGGATTTCAGCAGGTCTCTTCTCTGACTATACCTCTCC 1850
QY 901 AGAAATTTTCTTTTGAAGTACTACTGGAATTAACATTTGATGGATGCTCTACAAGCC 960
Db 1851 AGAAATTTTCTTTTGAAGTACTACTGGAATTAACATTTGATGGATGCTCTACAAGCC 1910
QY 961 TCATGATCTACAGCTCGGAAAGAAATATCTCTGCTGTGTTTCATATATATGTTGGTCTCTCA 1020
Db 1911 TCATGATCTACAGCTCGGAAAGAAATATCTCTGCTGTGTTTCATATATATGTTGGTCTCTCA 1970
QY 1021 ----- 1020
Db 1971 GGTGCGATTGGTGAATAATCGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTTAGC 2030
QY 1021 ----- 1020
Db 2031 CTCTCTAGGTTATGTGGTTGTAGTGATAGACAAACAGGGGATCTGTCTACCCGAGGCTTAA 2090
QY 1021 ----- 1020
Db 2091 ATTTGAAGGCGCTTTAAATATAAATGGGTCAATAGAAATTGACGATCAGGTGGAAGG 2150
QY 1021 ----- 1020
Db 2151 ACTCCAAATCTAGCTTCTCGATATGATTTTATTGACTTAGATCGTGTGGGCATCCACGG 2210
QY 1021 ----- 1020
Db 2211 CTGGTCTATGGAGGATACCTCTCCCTGATGGCAATTAATGCAGAGGTGAGATATCTTCAG 2270
QY 1021 GGTGCTATTCCTGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGA 1080
Db 2271 GGTGCTATTCCTGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGA 2330
QY 1081 AGTTATATGGTCACTCCCTGACCAAGATGAACAGGCTATTACTTAGGATCTGTGCCAT 1140

```
Db 2331 ACCTTATATGGGTCACCTCGACAGAAATGAACAGGGCTATTACTTAGGATCTGTGGCCAT 2390
Qy 1141 GCAAGCAGAAAAGTTCCTCTCTGAACCAAATCGTTTACTGCTCTTACATGGTTTCTCTGGA 1200
Db 2391 GCAAGCAGAAAAGTTCCTCTCTGAACCAAATCGTTTACTGCTCTTACATGGTTTCTCTGGA 2450
Qy 1201 TGAGAAATGTCATTTTGGCAATACACAGATATATTACTGAGTTTCTTGTAGGGGCTGAAA 1260
Db 2451 TGAGAAATGTCATTTTGGCAATACACAGATATATTACTGAGTTTCTTGTAGGGGCTGAAA 2510
Qy 1261 GCCATATGATTTACAGATCTATCTCTCAGGAGACACAGCATAAGAGTTCCTGAATCGGG 1320
Db 2511 GCCATATGATTTACAGATCTATCTCTCAGGAGACACAGCATAAGAGTTCCTGAATCGGG 2570
Qy 1321 AGACATATTAGTATGATCTCTTTTGCACTACCTTCAAGAAAACCTTGGATCAGTATTGC 1380
Db 2571 AGACATATTAGTATGATCTCTTTTGCACTACCTTCAAGAAAACCTTGGATCAGTATTGC 2630
Qy 1381 TGCTCTAAAAGTCATATA 1399
Db 2631 TGCTCTAAAAGTCATATA 2649
```

```
RESULT 3
US-09-925-065A-669313/G
; Sequence 669313, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 669313
; LENGTH: 1346
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-669313
```

```
Query Match 7.1%; Score 118; DB 7; Length 1346;
Best Local Similarity 100.0%; Pred. No. 4.3e-16;
Matches 118; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 873 CAGTCCTCTCTCTGACTATCTCTCCAGAAATTTCTCTTTGAAAGTACTACTGGAT 932
Db 118 CAGTCCTCTCTCTGACTATCTCTCCAGAAATTTCTCTTTGAAAGTACTACTGGAT 59
Qy 933 TTACATGTATGGGATGCTCTACAGCCTCATGATCTACAGCTGGAAGAAATATCC 990
Db 58 TTACATGTATGGGATGCTCTACAGCCTCATGATCTACAGCTGGAAGAAATATCC 1
```

```
RESULT 4
US-09-925-065A-743558/G
; Sequence 743558, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
```

```
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 743558
; LENGTH: 609
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-743558

Query Match 6.8%; Score 113.6; DB 7; Length 609;
Best Local Similarity 96.7%; Pred. No. 3.2e-15;
Matches 116; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1159 CTCGGAACCAAAATCGTTTACTGCTCTTACATGGTTTCTCGATGAGAATGTCCATTTTGC 1218
Db 205 CTCGAGACCAAAATCGTTTACTGCTCTTACATGGTTTCTCGATGAGAATGTCCATTTTGC 146
Qy 1219 ACATACCAGTATATTACTGAGTTTCTAGTAGGGCTGAAAGCCATATGATTACAGAT 1278
Db 145 ACATACCAGTATATTACTGAGTTTCTAGTAGGGCTGAAAGCCATATGATTACAGAT 86
```

```
RESULT 5
US-09-925-065A-743559/G
; Sequence 743559, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 743559
; LENGTH: 609
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-743559
```

```
Query Match 6.8%; Score 113.6; DB 7; Length 609;
Best Local Similarity 96.7%; Pred. No. 3.2e-15;
Matches 116; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1159 CTCGGAACCAAAATCGTTTACTGCTCTTACATGGTTTCTCGATGAGAATGTCCATTTTGC 1218
Db 205 CTCGAGACCAAAATCGTTTACTGCTCTTACATGGTTTCTCGATGAGAATGTCCATTTTGC 146
Qy 1219 ACATACCAGTATATTACTGAGTTTCTAGTAGGGCTGAAAGCCATATGATTACAGAT 1278
```

Db	145	ACATACAGTATTACTAGGTTTTTTAGTGAGGCTGGAAAGCCATGATTATTACAGGT	86
RESULT 6			
US-09-925-065A-818446/c			
; Sequence 818446, Application US/09925065A			
; Publication No. US20040181048A1			
; GENERAL INFORMATION:			
; APPLICANT: Wang, David G.			
; TITLE OF INVENTION: Identification and Mapping of Single			
; Nucleotide Polymorphisms in the Human Genome			
; FILE REFERENCE: 108827.135			
; CURRENT APPLICATION NUMBER: US/09/925,065A			
; PRIOR FILING DATE: 2001-08-08			
; PRIOR APPLICATION NUMBER: US 60/243,096			
; PRIOR FILING DATE: 2000-10-24			
; PRIOR APPLICATION NUMBER: US 60/252,147			
; PRIOR FILING DATE: 2000-11-20			
; PRIOR APPLICATION NUMBER: US 60/250,092			
; PRIOR FILING DATE: 2000-11-30			
; PRIOR APPLICATION NUMBER: US 60/261,766			
; PRIOR FILING DATE: 2001-01-16			
; PRIOR APPLICATION NUMBER: US 60/289,846			
; PRIOR FILING DATE: 2001-05-09			
; NUMBER OF SEQ ID NOS: 957086			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 818446			
; LENGTH: 612			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
US-09-925-065A-818446			
Query Match			
Best Local Similarity 4.4%; Score 73.4; DB 7; Length 612;			
Matches 74; Conservative 0; Mismatches 1; Indels 0; Gaps 0;			
Qy	1204	GAATGTCCATTTTGCACATACCATGATATTACTGAGTTTTTTAGTGAGGGCTGGAAAGCC	1263
Db	612	GAATGTCCATTTTGCACATACCATGATATTACTGAGTTTTTTAGTGAGGGCTGGAAAGCC	553
Qy	1264	ATATGATTACAGAT	1278
Db	552	ATATGATTACAGGT	538
RESULT 7			
US-11-079-463-1186			
; Sequence 1186, Application US/11079463			
; Publication No. US20060073161A1			
; GENERAL INFORMATION:			
; APPLICANT: Gary L. Breton			
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTEROIDES FR			
; FOR DIAGNOSTICS AND THERAPEUTICS			
; FILE REFERENCE: PAT000-03DIV2			
; CURRENT APPLICATION NUMBER: US/11/079,463			
; CURRENT FILING DATE: 2005-03-14			
; PRIOR APPLICATION NUMBER: US 60/128,705			
; PRIOR FILING DATE: 1999-04-09			
; PRIOR APPLICATION NUMBER: US 09/540,209			
; PRIOR FILING DATE: 2000-04-04			
; NUMBER OF SEQ ID NOS: 10444			
; SEQ ID NO 1186			
; LENGTH: 2238			
; TYPE: DNA			
; ORGANISM: B.fragilis			
US-11-079-463-1186			
Query Match			
Best Local Similarity 3.1%; Score 51.6; DB 18; Length 2238;			
Matches 166; Conservative 0; Mismatches 159; Indels 9; Gaps 1;			
Qy	1029	TTGTGGGGCCCCCACTCTGTGGATCTTCTATGATACAGGATACACGGAACGTTATA	1088

Db	1907	TTGCTGTGGCCGACCTTACAGACTGGAAATATTACGATACAGTATATACCGAAGCTTTA	1966
Qy	1089	TGGGTACACCTGACAGAAATGAACAGGCTATTACTTAGGATCTGTGGCATGCAAGCAG	1148
Db	1967	TGCGCAGCCGAAAGAAATGCCGAGGCTATAAAGCAGCTTCAGCATTCAGCCGTGAG	2026
Qy	1149	AAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGTTTCTCGATGAGAAATG	1208
Db	2027	ACAACTTGCATGGTAAACC-----TGCTCTTGTACACGTTATGCGAGATGATAATG	2077
Qy	1209	TCCATTTTGCACATACCATGATATTACTGAGTTTTTTAGTGAGGCTGGAAGCCATATG	1268
Db	2078	TTCACTTCCAGAACTGTACAGAAATATGCAGAGCACTGTGTAACAATCGGAAAAACAGTT	2137
Qy	1269	ATTTACAGATCTATCTCTCAGGAGAGACACAGCATAAGAGTTCTCTGAATCGGGAGAACAT	1328
Db	2138	ATATGAGGTTATACACCAACCGGAATCATAGATCTATGTTGGAATACCCGTAACCACT	2197
Qy	1329	ATGAATGCACTCTTTTGGCACTACCTTCAAGAAAA	1362
Db	2198	TGTATACGAAGCTGACGAACCTCTTCCGGAATAA	2231
RESULT 8			
US-10-932-182A-81332			
; Sequence 81332, Application US/10932182A			
; Publication No. US20060046253A1			
; GENERAL INFORMATION:			
; APPLICANT: NAKAO, YOSHIHIRO			
; APPLICANT: NAKAMURA, NORIHIISA			
; APPLICANT: KODAMA, YUKIKO			
; APPLICANT: FUJIMURA, TOMOKO			
; APPLICANT: ASHIKARI, TOSHIHIKO			
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS			
; FILE REFERENCE: 030685-043			
; CURRENT APPLICATION NUMBER: US/10/932,182A			
; CURRENT FILING DATE: 2004-09-02			
; NUMBER OF SEQ ID NOS: 197023			
; SOFTWARE: PatentIn version 3.3			
; SEQ ID NO 81332			
; LENGTH: 378			
; TYPE: DNA			
; ORGANISM: Saccharomyces pastorianus			
US-10-932-182A-81332			
Query Match			
Best Local Similarity 3.0%; Score 50.6; DB 11; Length 378;			
Matches 117; Conservative 0; Mismatches 89; Indels 3; Gaps 1;			
Qy	1028	ATTGCTGGGGCCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAACGTTAT	1087
Db	1	ATGCTCTCGGCGCAGTAACTGACATTTGATGATTCGTTTACACGGAAGATAT	60
Qy	1088	ATGGGTACCCCTGACAGAAATGAACAGGCTATTACTTAGGATCTGTGGCATGCAAGCA	1147
Db	61	ATGAACCAACCATCGGAAAAACCATGAAGCTATTTTGAAGTATCCA---CCATCCAAAT	117
Qy	1148	GAAGAATTCCTCTGNAACCAATCGTTTACTGCTTACATGCTTCTCGATGAGAT	1207
Db	118	TTCAAAATCATTCGAATCTTTTAAAGCGTCTTTTCATTTGTGACGGAACCTTTTGTATGAT	177
Qy	1208	GTCCATTTTGCACATACCATGATATTACT	1236
Db	178	GTTACATACAAATACTTTTAGATTAGT	206
RESULT 9			
US-10-932-182A-81332			
; Sequence 81332, Application US/10932182A			
; Publication No. US20060046253A1			
; GENERAL INFORMATION:			
; APPLICANT: NAKAO, YOSHIHIRO			

```
; APPLICANT: NAKAMURA, NORIHIISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: 030685-043
; CURRENT APPLICATION NUMBER: US/10/932,182A
; CURRENT FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 81332
; LENGTH: 378
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-10-932-182A-81332

Query Match      3.0%; Score 50.6; DB 11; Length 378;
Best Local Similarity 56.8%; Pred. No. 0.23;
Matches 117; Conservative 0; Mismatches 89; Indels 3; Gaps 1;

Qy 1028 ATTGCTGGGGCCCACTCTGTGGATCTTTCTATGATACAGGATACAGGAACGTTAT 1087
Db 1 ATGGCTGTGCGCGCAGTAACAACCTGGACATTTGATGATCCGTTTACCGGAAGATAT 60
Qy 1088 ATGGCTCACCTGACCAAGATGAACAGGGCTATTACTTTAGGATCTGTGGCCATGCAAGCA 1147
Db 61 ATGAACCAACATCGGAAACCACTGAAGGCTATTTTGAAGTATCCA---CCATCCAAAT 117
Qy 1148 GAAAGATTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGCTTTCTGGATGAGAT 1207
Db 118 TTCAATCATTCGNAATCTTTAAAGCGTCTTTTCATTGTGACGGAACCTTTTCGATGATAT 177
Qy 1208 GTCCATTTTGACATACCAAGTATATTACT 1236
Db 178 GTTCACATACAAATACTTTTAGATTAGT 206
```

```
RESULT 10
US-10-932-182A-5649
; Sequence 5649, Application US/10932182A
; Publication No. US20060046253A1
; GENERAL INFORMATION:
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHIISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: 030685-043
; CURRENT APPLICATION NUMBER: US/10/932,182A
; CURRENT FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 5649
; LENGTH: 2778
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-10-932-182A-5649

Query Match      3.0%; Score 49.6; DB 11; Length 2778;
Best Local Similarity 55.8%; Pred. No. 0.65;
Matches 116; Conservative 0; Mismatches 89; Indels 3; Gaps 1;

Qy 1026 CTATTGCTGGGGCCCGAGTCACTCTGTGGATCTTCTATGATACAGGATACAGGAACGTT 1085
Db 2414 CAATGGCTGTGACCAAGTAAACAATCGACACTGTATGATCTTACCTACACCGAAGGT 2473
Qy 1086 ATATGGGTCACTGACCAAGATGAACAGGGCTATTAGGATCTGTGGCCATGCAAG 1145
Db 2474 ATATGAATCAACCATCGGAATGAAGGGTATTTTGACATATCTA---CTATTAAA 2530
Qy 1146 CAGAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGGTTTCTCGATGAGA 1205
```

```
Db 2531 ATTACAAAGTCTTTCGATCGCTGAACGGAATCTTTGTTATGACGGGACCTCTGATGATA 2590
Qy 1206 ATGTCCTATTTCACATACCAAGTATATT 1233
Db 2591 ATGTTCAATATCAAAATACGTTTAGATT 2618

RESULT 11
US-10-932-182A-5649
; Sequence 5649, Application US/10932182A
; Publication No. US20060046253A1
; GENERAL INFORMATION:
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHIISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: 030685-043
; CURRENT APPLICATION NUMBER: US/10/932,182A
; CURRENT FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 5649
; LENGTH: 2778
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-10-932-182A-5649

Query Match      3.0%; Score 49.6; DB 11; Length 2778;
Best Local Similarity 55.8%; Pred. No. 0.65;
Matches 116; Conservative 0; Mismatches 89; Indels 3; Gaps 1;

Qy 1026 CTATTGCTGGGGCCCGAGTCACTCTGTGGATCTTCTATGATACAGGATACAGGAACGTT 1085
Db 2414 CAATGGCTGTGACCAAGTAAACAATCGACACTGTATGATCTTACCTACACCGAAGGT 2473
Qy 1086 ATATGGGTCACTGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAG 1145
Db 2474 ATATGAATCAACCATCGGAATGAAGGGTATTTTGACATATCTA---CTATTAAA 2530
Qy 1146 CAGAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGGTTTCTCGATGAGA 1205
Db 2531 ATTACAAAGTCTTTCGAAATCGCTGAACGGAATCTTTGTTATGACGGGACCTCTGATGATA 2590
Qy 1206 ATGTCCTATTTCACATACCAAGTATATT 1233
Db 2591 ATGTTCAATATCAAAATACGTTTAGATT 2618

RESULT 12
US-11-208-288-1
; Sequence 1, Application US/11208288
; Publication No. US20060051366A1
; GENERAL INFORMATION:
; APPLICANT: CHANG, Chiwen
; TITLE OF INVENTION: USE OF SOLUBLE CD26 AS INHIBITOR OF
; FILE REFERENCE: 39533-0001
; CURRENT APPLICATION NUMBER: US/11/208,288
; CURRENT FILING DATE: 2005-08-18
; PRIOR APPLICATION NUMBER: US 60/605,013
; PRIOR FILING DATE: 2004-08-26
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3332
; TYPE: DNA
; ORGANISM: homo sapiens
US-11-208-288-1

Query Match      2.7%; Score 45.6; DB 18; Length 3332;
Best Local Similarity 46.1%; Pred. No. 5.2;
```


Matches 188; Conservative 0; Mismatches 219; Indels 1; Gaps 1;

QY 1020 AGTTGCTATTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGG 1079

Db 1943 AGTGTGGAAATAGCCGTGGCGCTGTATCCGGTGGGAGTACTATGACTCAGTGTACACAG 2002

QY 1080 AACGTTATATGGTCCACCCTGACCAGAAATGAACAGGGCTATTACTTAGGATCTGGGCA 1139

Db 2003 AACGTTATATGGTCTCCCACTCCAGAGACAACCTTGACCAATTACAGAAATTCACACAG 2062

QY 1140 TCGAAGCAGAAAGTTCCCTCTCTGAAACAAATCGTTTACTGCTCTTACATGATGTTCCCTGG 1199

Db 2063 TCATGAGCAGAGCTGAAAATTTAAACAAGTTGAGTACCTCTTATTCATGGAACAGCAG 2122

QY 1200 ATGAGAATGTCATTTTGCATACATACAGTATATATCTAGTGTATTTTGTAGGCGGTGAA 1259

Db 2123 ATGATAACGTTTCACTTTCCAGCAGTCAGTCCAGATCTCCAAAGCCCTGGTGCATGTGGAG 2182

QY 1260 AGCCATATGATTTACAGATCTATCTCAGGAGAGACACAGATAGAGTTCCTGAATCGG 1319

Db 2183 TGGATTTCCAGCAATGTGGTATATCTGATGAGACCATGGAATAGCTAGCAGCAGCAC 2242

QY 1320 GAGAACATTATGAACATGCTTTTGGCACTACCTTCAAGAAAAACCTTGGATCAAGTATTG 1379

Db 2243 ACCACATATATATACCCACATGAGCCACTTCATAAAAACAATGTTTCTCTTTACCT-TAG 2301

QY 1380 CTGCTCTAAAGTGATATAATTTTACCTGTGTAGAACTCTCTGGTAT 1427

Db 2302 CACCTCAAATACCATGCAATTTAAAGCTTATTAAACTCATTTTGT 2349

RESULT 13

US-10-501-035-34

; Sequence 34, Application US/10501035

; Publication No. US20060046249A1

; GENERAL INFORMATION:

; APPLICANT: Bristol-Myers Squibb Company

; TITLE OF INVENTION: IDENTIFICATION OF POLYNUCLEOTIDES AND POLYPEPTIDE FOR PREDICTING

; TITLE OF INVENTION: ACTIVITY OF COMPOUNDS THAT INTERACT WITH PROTEIN TYROSINE KINASE

; TITLE OF INVENTION: AND/OR PROTEIN TYROSINE KINASE PATHWAYS

; FILE REFERENCE: D0185 PCT

; CURRENT APPLICATION NUMBER: US/10/501,035

; PRIOR FILING DATE: 2004-07-09

; PRIOR APPLICATION NUMBER: US 60/350,061

; PRIOR FILING DATE: 2002-01-18

; NUMBER OF SEQ ID NOS: 795

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 34

; LENGTH: 3407

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-501-035-34

Query Match 2.7%; Score 45.6; DB 11; Length 3407;

Best Local Similarity 46.1%; Pred. No. 5.3;

Matches 188; Conservative 0; Mismatches 219; Indels 1; Gaps 1;

QY 1020 AGTTGCTATTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGG 1079

Db 2018 AGTGTGAATAGCGTGGCGCTGTATCCCGTGGGAGTACTATGACTCAGTGTACACAG 2077

QY 1080 AACGTTATATGGTCCACCCTGACCAGAAATGAACAGGGCTATTACTTAGGATCTGTGGCCA 1139

Db 2078 AACGTTATATGGTCTCCCACTCCAGAGACAACCTTGACCAATTACAGAAATTCACACAG 2137

QY 1140 TCGAAGCAGAAAGTTCCCTCTCTGAAACAAATCGTTTACTGCTCTTACATGTTTCTCTGG 1199

Db 2138 TCATGAGCAGAGCTGAAAATTTTAAACAAGTTGAGTACCTCTTATTCATGGAACAGCAG 2197

QY 1200 ATGAGAATGTCATTTTGCATACATACAGTATATTAAGTGTATTTTGTAGTGGCGGTGAA 1259

Db 2198 ATGATAACGTTTCACTTTACAGCAGTCAGTCCAGATCTCCAAAGCCCTGGTGCATGTGGAG 2257

QY 1260 AGCCATATGATTTACAGATCTATCTCAGGAGAGACACAGATAGAGTTCCTGAATCGG 1319

Db 2258 TGGATTTCCAGGCAATGTGGTATATCTGATGAAGACCATGGAATAGCTAGCAGCAGCAC 2317

QY 1320 GAGAACATTATGAACATGCTTTTGGCACTACCTTCAAGAAAAACCTTGGATCAAGTATTG 1379

Db 2318 ACCACATATATATACCCACATGAGCCACTTCATAAAAACAATGTTTCTCTTTACCT-TAG 2376

QY 1380 CTGCTCTAAAGTGATATAATTTTACCTGTGTAGAACTCTCTGGTAT 1427

Db 2377 CACCTCAAATACCATGCAATTTAAAGCTTATTAAACTCATTTTGT 2424

RESULT 14

US-10-131-826A-181

; Sequence 181, Application US/10131826A

; Publication No. US20050245730A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Beresini, Maureen

; APPLICANT: DeForge, Laura

; APPLICANT: Desnoyers, Luc

; APPLICANT: Filvaroff, Ellen

; APPLICANT: Gao, Wei-Qiang

; APPLICANT: Gerritsen, Mary E.

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Sherwood, Steven

; APPLICANT: Smith, Victoria

; APPLICANT: Stewart, Timothy A.

; APPLICANT: Tamas, Daniel

; APPLICANT: Watanabe, Colin K

; APPLICANT: Wood, William

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; TITLE OF INVENTION: ACIDS ENCODING THE SAME

; FILE REFERENCE: P3330RJC128

; CURRENT APPLICATION NUMBER: US/10/131,826A

; CURRENT FILING DATE: 2002-04-24

; PRIOR APPLICATION NUMBER: 60/049911

; PRIOR FILING DATE: 1997-06-18

; PRIOR APPLICATION NUMBER: 60/056974

; PRIOR FILING DATE: 1997-08-26

; PRIOR APPLICATION NUMBER: 60/059113

; PRIOR FILING DATE: 1997-09-17

; PRIOR APPLICATION NUMBER: 60/059115

; PRIOR FILING DATE: 1997-09-17

; PRIOR APPLICATION NUMBER: 60/059117

; PRIOR FILING DATE: 1997-09-17

; PRIOR APPLICATION NUMBER: 60/059122

; PRIOR FILING DATE: 1997-09-17

; PRIOR APPLICATION NUMBER: 60/059184

; PRIOR FILING DATE: 1997-09-17

; PRIOR APPLICATION NUMBER: 60/059263

; PRIOR FILING DATE: 1997-09-18

; PRIOR APPLICATION NUMBER: 60/059352

; PRIOR FILING DATE: 1997-09-19

; PRIOR APPLICATION NUMBER: 60/059588

; PRIOR FILING DATE: 1997-09-19

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 550

; SEQ ID NO 181

; LENGTH: 1624

; TYPE: DNA

; ORGANISM: Homo Sapien

; FEATURE:

; NAME/KEY: unsure

; LOCATION: 1560-1561

; OTHER INFORMATION: unknown base

US-10-131-826A-181

Query Match 2.7%; Score 44.8; DB 10; Length 1624;

```
Best Local Similarity 63.2%; Pred. No. 6.5;
Matches 67; Conservative 0; Mismatches 39; Indels 0; Gaps 0;
QY 1564 AGTAATCTAATACCTTAACCCACATGCTCAAAATCAATGATACATATTCCTGAGAC 1623
DB 1514 AGAAATTAATGTTGAATCCAAATCCTAGAAATGAATGAGTTTNNTTCCAAAAA 1573
QY 1624 CCAGCAATACCAATGAATTAATTAATAAATAAATAAATAAATAAATAA 1669
DB 1574 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1619
```

```
RESULT 15
US-10-973-115B-181
; Sequence 181, Application US/10973115B
; Publication No US20060040351A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Desrogers, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiung
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC ACIDS ENCODING
; FILE REFERENCE: 39870-3330R1C300C1
; CURRENT APPLICATION NUMBER: US/10/973,115B
; CURRENT FILING DATE: 2004-10-22
; PRIOR APPLICATION NUMBER: US 10/145,747
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: US 10/028,072
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: PCT/US00/32678
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: US 09/581,742
; PRIOR FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: PCT/US00/05746
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 60/135,736
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 60/123,090
; PRIOR FILING DATE: 1999-03-05
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 181
; LENGTH: 1624
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 1560-1561
; OTHER INFORMATION: unknown base
US-10-973-115B-181
```

```
Query Match 2.7%; Score 44.8; DB 11; Length 1624;
Best Local Similarity 63.2%; Pred. No. 6.5;
Matches 67; Conservative 0; Mismatches 39; Indels 0; Gaps 0;
QY 1564 AGTAATCTAATACCTTAACCCACATGCTCAAAATCAATGATACATATTCCTGAGAC 1623
DB 1514 AGAAATTAATGTTGAATCCAAATCCTAGAAATGAATGAGTTTNNTTCCAAAAA 1573
QY 1624 CCAGCAATACCAATGAATTAATTAATAAATAAATAAATAAATAAATAA 1669
```

```
DB 1574 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1619
Search completed: May 4, 2006, 09:26:05
Job time : 2285.71 secs
```

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus.n2p model

Run on: May 2, 2006, 00:41:48 ; Search time 50.6744 Seconds
(without alignments)
2894.257 Million cell updates/sec

Title: US-10-825-632-6
Perfect score: 2957
Sequence: 1 aacaggtacagcaaatccta.....aaaaaaaaaaaaaaaaaa 1669

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2443163 seqs, 439378781 residues
Total number of hits satisfying chosen parameters: 4886326

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ n2p.model -DEV=xlp
-Q=/abss/ABSSWEB_spool/US10825632/runat_01052006_105944_3214/app_query.fasta_1
-DB=A Geneseq -QFMT=fastan -SUFFIX=n2p.rag -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTPMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HCST=abs02p
-USER=US10825632 @CGN_1_1_605 @runat_01052006_105944_3214 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORE=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A Geneseq 21.*
1: Geneseq1980s.*
2: Geneseq1990s.*
3: Geneseq2000s.*
4: Geneseq2001s.*
5: Geneseq2002s.*
6: Geneseq2003as.*
7: Geneseq2003bs.*
8: Geneseq2004s.*
9: Geneseq2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2482	83.9	465	AA847189	Human DPP
2	2482	83.9	724	AB97362	Novel hum
3	2482	83.9	782	AB97361	Novel hum
4	2447.5	82.8	831	ABU92026	Human pro
5	2441.5	82.6	632	AA93565	Human pro
6	2431.5	82.2	738	ADT04044	Human pro
7	2422	81.9	824	ABU92030	Human pro
8	2422	81.9	882	AA847187	Human DPP
9	2422	81.9	882	ABG61591	Human DPP

10	2422	81.9	882	5	AAG78415	Aag78415 Amino aci
11	2422	81.9	882	5	AAE24170	Aae24170 Human dip
12	2422	81.9	882	5	AAU74749	Aau74749 Human pro
13	2422	81.9	882	5	AD117086	Adi17086 Human NOV
14	2422	81.9	882	6	ABU07720	Abu07720 Human ser
15	2411	81.5	587	5	ADR41398	Adr41398 Human CD-
16	2326	78.7	883	5	AD117085	Adi17085 Murine NO
17	2244.5	75.9	580	5	AAE14337	Aae14337 Human pro
18	2134.5	72.2	516	6	ABU92029	Abu92029 Human pro
19	1992	67.4	746	8	AD116334	Adi16334 Human pro
20	1957.5	66.2	690	5	ABG61594	Abg61594 Human DPR
21	1915	64.8	493	7	ADE78977	Ade78977 Human pro
22	1808	61.1	658	5	ABG61600	Abg61600 Human DPR
23	1808	61.1	661	5	ABG61596	Abg61596 Human DPR
24	1540	52.1	613	5	ABG61601	Abg61601 Human DPR
25	1494.5	50.5	689	5	AD117084	Adi17084 Human NOV
26	1494.5	50.5	755	8	ADQ67811	Adq67811 Novel hum
27	1494.5	50.5	830	5	AAE24171	Aae24171 Human dip
28	1494.5	50.5	863	5	ABG61592	Abg61592 Human DPP
29	1494.5	50.5	863	5	AD117083	Adi17083 Human NOV
30	1494.5	50.5	892	5	ABG61602	Abg61602 Human DPR
31	1494.5	50.5	892	5	ABG61604	Abg61604 Human DPR
32	1494.5	50.5	892	5	AB898134	Ab898134 Human PMM
33	1494.5	50.5	892	8	ADS10951	Ads10951 Human the
34	1494.5	50.5	969	5	AAE24168	Aae24168 Human dip
35	1490.5	50.4	581	5	ADR41521	Adr41521 Human CD-
36	1488.5	50.3	869	5	AAE24169	Aae24169 Alternati
37	1486.5	50.3	863	5	AD116690	Adi16690 Human NOV
38	1486.5	50.3	863	5	AD116688	Adi16688 Human NOV
39	1486.5	50.3	863	8	ADN42344	Adn42344 Human NOV
40	1474.5	49.9	863	8	ADN42342	Adn42342 Human nov
41	1445	48.9	879	5	ABG61607	Abg61607 Human DPR
42	1445	48.9	879	5	ABG61608	Abg61608 Human DPR
43	1445	48.9	879	8	ADS10953	Ads10953 Human the
44	1412	47.8	597	7	ADE808485	Ade808485 Novel pro
45	1377.5	46.6	310	4	AA847188	Aab47188 Human DPP

ALIGNMENTS

RESULT 1
AA847189
ID AA847189 standard; protein; 465 AA.
XX
AC AA847189;
XX
DT 29-JUN-2001 (first entry)
XX
DE Human DPP8 318Thr-658Val+759Ala-882Ile.
XX
KW Human; dipeptidyl aminopeptidase; DPP8; prolol oligopeptidase;
KW dipeptidyl peptidase; DPPIV; T cell; cleavage; diarrhoea;
KW growth hormone deficiency; glucose level; mucosal regeneration;
KW non-insulin dependent diabetes mellitus; glucose intolerance;
KW immunosuppression.
XX
OS Homo sapiens.
XX
PN WO200119866-A1.
XX
PD 22-MAR-2001.
XX
PF 11-SEP-2000; 2000WO-AU001085.
XX
PR 10-SEP-1999; 99AU-00002762.
PR 18-FEB-2000; 2000AU-00005709.
XX
PA (UNSY) UNIV SYDNEY.
XX
PI Abbott CA, Gorell MD;
XX WPI; 2001-281520/29.
DR N-PSDB; AAC85696.

XX New human dipeptidyl aminopeptidase (DPP8) useful for cleaving
 PT substrates, identifying inhibitors of DPP8 catalytic activity which have
 PT therapeutic uses, and for detecting activated T cells.

PS Claim 5; Page 71-72; 78pp; English.

XX The sequences given in ABA47188-90 represent fragments of human
 CC dipeptidyl aminopeptidase (DPP8). DPP8 has substrate specificity for H-
 CC Ala-Pro-pNA, H-Gly-Pro-pNA and H-Arg-Pro-pNA. Therefore, it is a prolyl
 CC oligopeptidase and a dipeptidyl peptidase, because it is capable of
 CC hydrolysing the peptide bond C-terminal to Pro in each of these
 CC compounds. DPP8 is homologous with human DPP4. DPP8 is useful for
 CC cleaving a substrate, and for detecting an activated T cell which
 CC involves measuring the level of DPP8 gene expression in a T cell. The
 CC level of DPP8 expression is detected by detecting the amount of DPP8 RNA
 CC in the cell. It is also useful for identifying a molecule capable of
 CC inhibiting the cleavage of the substrate by DPP8. Molecules identified as
 CC inhibiting DPP8 catalytic activity may be useful for treating diarrhoea,
 CC growth hormone deficiency, lowering glucose levels in non-insulin
 CC dependent diabetes mellitus and other disorders involving glucose
 CC intolerance, enhancing mucosal regeneration and as immunosuppressants

XX Sequence 465 AA;

Alignment Scores:
 Pred. No.: 1,21e-254 Length: 465
 Score: 2482.00 Matches: 465
 Percent Similarity: 100.0% Conservatives: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 83.9% Indels: 0
 DB: 4 Gaps: 0

US-10-825-632-6 (1-1669) x ABA47189 (1-465)

Qy 2 ACAGGTACGAACTCTAAAGTCACTTTTAAGATGTCAGAAATAATGATGTCTGAA 61
 Db 1 ThrGlyThrAlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGlu 20
 Qy 62 GGAAGGATCATAGATGTCATAGATAAGAACTAATTCACACCTTTTGAGATTCATTTGAA 121
 Db 21 GlyArgIleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGlu 40
 Qy 122 GGAGTTGAATATATGCCAGACTGGATGGACTCTCGAGGAAATAATGCTGGTCCATC 181
 Db 41 GlyValGluLysIleAlaArgAlaGlyTrpThrProGluGlyLysIleAlaTrpSerIle 60
 Qy 182 CTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAATATTATC 241
 Db 61 LeuLeuAspArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIle 80
 Qy 242 CCAGTAGAAGATGATGTTATGAAAGGCAGAGACTCATTGAGTCAGTGCCTGATCTGTG 301
 Db 81 ProValGluAspValMetGluArgGlnArgLeuIleGluSerValProAspSerVal 100
 Qy 302 ACGCCACTAATTATCTATGAGAAACACACACATCTGGATATATATCCATGACATCTTT 361
 Db 101 ThrProLeuIleIleLysGluIleThrAspIleTrpIleAsnIleHisAspIlePhe 120
 Qy 362 CANGTTTTCCCAAGTCACAGAGAAATTCAGTTTATTTTGGCTCTGATCCAA 421
 Db 121 HisValPheProGlnSerHisGluGluIleGluPheIlePheAlaSerGluCysLys 140
 Qy 422 ACAGGTTTCGGTCATTTATACAAATTCATCTATTATTAAGAAAGCAATAATAACGA 481
 Db 141 ThrGlyPheArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysIleArg 160
 Qy 482 TCAGTGGTGGGCTCGCTCCAAAGTGATTTCAAGTGCTCTATCAAGAGGAGATAGCA 541
 Db 161 SerSerGlyGlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAla 180
 Qy 542 ATTACAGTGGTGAATGGAAAGTCTTGGCCGGCATGGATCTAATATCCAAAGTTGATGAA 601

Db IleThrSerGlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGlu 200
 Qy 602 GTCAGAAAGGTGTATATATTTGAAGGCACCAAAAGACTCCCTTTAGACATCATCCTGTAC 661
 Db 201 ValArgArgLeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeuTyr 220
 Qy 662 GTAGTCAGTTACGTAATCTCGAGAGGTGACAAGGCTGACTGACCCGTGGCTACTACAT 721
 Db 221 ValValSerTyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHis 240
 Qy 722 TCTTGTGTCATCAGTCAGACTGTGACTTCTTTATAAGTAAGTAAAGTAAACACCAAGAAT 781
 Db 241 SerCysCysIleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsn 260
 Qy 782 CCACACTGTGTGCTTTCACAGCTATCAAGCTCTCAAGATGATGACCCCAACTTGCACAAACA 841
 Db 261 ProHisCysValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThr 280
 Qy 842 AAGGAATTTTGGGCCACCATTTTGGATTCAGCAGGTCTCTTCTCTGACTATCTCTCCA 901
 Db 281 LysGluPheTrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProPro 300
 Qy 902 GAAATTTCTCTTTGAAAGTACTACTGGATTACATTTACATTTGATGGATGCTCTACAAGCT 961
 Db 301 GluIlePheSerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysPro 320
 Qy 962 CATGATCTACAGCTCGAAAGAAATATCTACTGTCTGTCTCATATATATGTTGCTCTCAG 1021
 Db 321 HisAspLeuGlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyProGln 340
 Qy 1022 GTTGCTATTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAA 1081
 Db 341 ValAlaIleAlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGlu 360
 Qy 1082 CGTTATATGGTCACTCCCTGACCAAGTGAACAGGGCTATTACTTAGGATCTGTGGCCATG 1141
 Db 361 ArgTyrMetGlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMet 380
 Qy 1142 CAAGCAGAAAGTCCCTCTCGAACCAAACTGTTTACTGCTCTTACATGTTTCTCTGGAT 1201
 Db 381 GlnAlaGluLysPheProSerGluProAsnArgLeuLeuLeuLeuHisGlyPheLeuAsp 400
 Qy 1202 GAGAATGTCATTTTGCACATACCAAGTATATTTACTAGTATTTTGTAGTGAGGCTGGAAAG 1261
 Db 401 GluAsnValHisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLys 420
 Qy 1262 CCATATGATTTACAGATCTATCTCAGGAGAGACACAGCATAAGAGTTCTTGAATCGGA 1321
 Db 421 ProTyrAspLeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGly 440
 Qy 1322 GAACATTTGAATCTGCTCTTTGCACTACCTTCAAGAAACCTTGGATCACGTTATGCT 1381
 Db 441 GluHisTyrGluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAla 460
 Qy 1382 GCTCTAAAGTGATA 1396
 Db 461 AlaLeuLysValIle 465

RESULT 2
 ABB97362
 ID ABB97362 standard; protein; 724 AA.
 XX
 AC ABB97362;
 XX
 DT 27-JUN-2002 (first entry)
 XX
 DE Novel human protein SEQ ID NO: 630.
 XX
 KW Human; antianaemic; vulnery; antiinflammatory; immunomodulator;
 KW antiinfertility; cerebroprotective; cytostatic; rheumatic; gene therapy;
 KW neuroprotective; antiparkinsonian; protein therapy; EST;
 KW expressed sequence tag.

AC ABB97361;
 DT 27-JUN-2002. (first entry)
 XX
 DE Novel human protein SEQ ID NO: 629.
 XX
 KW Human; antinaemic; vulnery; antinflammatory; immunomodulator;
 XX antifertility; cerebroprotective; cytostatic; rheumatic; gene therapy;
 KW neuroprotective; antiparkinsonian; protein therapy; EST;
 XX expressed sequence tag.
 XX
 OS Homo sapiens.
 XX
 FN WO200222660-A2.
 XX
 PD 21-MAR-2002.
 XX
 PF 10-SEP-2001; 2001WO-US026015.
 XX
 PR 11-SEP-2000; 2000US-00659671.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
 PI Xue AJ, Yang Y, Wehrman T, Drmanac RT;
 XX
 DR WPT; 2002-292408/33.
 DR N-PSDB; ABN32547.
 XX

XX An isolated polynucleotide for treating diseases associated with its
 PT encoded polypeptide such as cancer and multiple sclerosis.
 XX
 PS Example 2; SEQ ID NO 629; 509pp; English.
 XX

XX The present invention provides the protein and coding sequences of 444
 CC novel human proteins. These were isolated from expressed sequences tags
 CC (ESTs). They can be used to stimulate cell growth, to regulate
 CC haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth
 CC e.g. in burn treatment, to regulate the immune system e.g. to treat
 CC multiple sclerosis, to regulate activin or inhibin e.g. to treat
 CC infertility, to regulate haemostasis or thrombolysis e.g. to treat stroke
 CC and cancer, to screen for drugs, to treat inflammatory conditions e.g.
 CC rheumatoid arthritis, and to treat nervous system disorders e.g.
 CC Parkinson's disease. The present sequence is a protein of the invention
 XX

XX Sequence 782 AA;

Alignment Scores:
 Pred. No.: 1-61e-254 Length: 782
 Score: 2482.00 Matches: 465
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 83.9% Indels: 0
 DB: 5 Gaps: 0

US-10-825-632-6 (1-1669) x ABB97361 (1-782)

QY 2 ACAGGTACAGCAATCCTTAAGTCACCTTTAAGATGTCAGAAATATGATGCTGAA 61
 DB 318 ThrGlyThrAlaAsnProLysValThrPheLysMetSerGluLeuMetIleAspAlaGlu 337
 QY 62 GGAAGATCATAGTGTCTATAGATAGAGAACTAATCAACTTTTGAGATTCATTGAA 121
 DB 338 GlyArgIleAspValIleAspLysGluLeuIleGlnProPheGluLeuPheGlu 357
 QY 122 GGAGTTGAATPATTGCCAGAGCTGGATGGACTCCTGAGGAAATATGCTGGTCCATC 181
 DB 358 GlyValGluTyrIleAlaArgAlaGlyTrpThrProGluGlyLysTrpAlaTrpSerIle 377
 QY 182 CTATAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAAATATTATTC 241
 DB 378 LeuLeuAspArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIle 397

QY 242 CCAGTACAGATCATGTTATGAAAGGCAGAGACTCATTGAGTCAGTCCCTGATTCTGTG 301
 DB 398 ProValGluAspValMetGluArgGlnArgLeuIleGluSerValProAspSerVal 417
 QY 302 ACGCCACTAATTATCTATGAAGAAACAACAGACATCTGGATAAATATCCATGACATCTTT 361
 DB 418 ThrProLeuIleIleTyrGluGluThrThrAspIleIleTrpIleAsnIleHisAspIlePhe 437
 QY 362 CATGTTTTCCCAAGTACAGAGAGAAATTTAGTTTATTTTTCCTCTGAAATGCAAA 421
 DB 438 HisValPheProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLys 457
 QY 422 ACAGTTTCCGTCATTTATACAAAATTACATCTATTTTAAAGGAAGCAAAATATAACGA 481
 DB 458 ThrGlyPheArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArg 477
 QY 482 TCCAGTGGTGGCTGCTCTCCAAAGTATTTCAAGTGTCTCTATCAAGAGAGATAGCA 541
 DB 478 SerSerGlyGlyLeuProAlaProSerAspPheLysCysProIleLysGluLysIleAla 497
 QY 542 ATTACCAAGTGTGAATGGGAAGTTCTTGGCCGGCATGGATCTATATCCAAAGTTGATGAA 601
 DB 498 IleThrSerGlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGlu 517
 QY 602 GTCAGAAGCTGTATATTTTGAAGGCACCAAAAGACTCCCTTTAGACATFCACCTGTAC 661
 DB 518 ValArgArgLeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyr 537
 QY 662 GTAGTCAGTTAGCTAAATCCTCGAGAGGTGACAAAGGCTGACTGACCGTGGCTACTCAT 721
 DB 538 ValValSerTyrValAsnProGlyValThrArgLeuThrAspArgGlyTyrSerHis 557
 QY 722 TCTTGTGTCATCAGTCAGTGTGACTTCTTTTATAAGTAAGTATAGTAAACCAAGAAAT 781
 DB 558 SerCysCysIleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsn 577
 QY 782 CCACACTGTGTGCTCCCTTTACAAAGCTCTCAAGTCTCTGAAGATGACCAACTTGCAGAA 841
 DB 578 ProHisCysValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThr 597
 QY 842 AAGCAATTTTGGGCCACCACTTTGGATTGACAGAGTCTCTCTCTGACTACTACTCTCCA 901
 DB 598 LysGluPheTrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProPro 617
 QY 902 GAAATTTTCTCTTTTGAAGTACTACTGGATTTACATTTGATGGATGCTCTCAAGCCT 961
 DB 618 GluIlePheSerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysPro 637
 QY 962 CATGATCTACAGCTGGAAAGAAATATCCTACTCTGCTGTCATATATGTTGCTCTCAG 1021
 DB 638 HisAspLeuGlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGln 657
 QY 1022 GTTGCTATTGCTGGGCCCCAGTCACTCTGTGGATCTCTTATGATACAGGATACAGGAA 1081
 DB 658 ValAlaIleAlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGlu 677
 QY 1082 CGTTATATGGGTACCCCTGACAGAAATGAAACAGGGCTATTACTTATGATCTGTGGCATG 1141
 DB 678 ArgTyrMetGlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMet 697
 QY 1142 CAAGCAGAAAGATTTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGTTTCTCGGAT 1201
 DB 698 GlnAlaGluLysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAsp 717
 QY 1202 GAGAATGTCCTATTTTGCATACACAGTATATTACTGAGTGTGTTTGTAGTGGGCTGGAAG 1261
 DB 718 GluAsnValIlePheAlaHisThrSerIleLeuLeuSerPheLeuValAlaGlyLys 737
 QY 1262 CCATATGATTACAGATCTATCTCCAGAGAGACACAGCATAGAGTTCTCGAATCGGGA 1321
 DB 738 ProTyrAspLeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGly 757
 QY 1322 GAACATTATGAACACTCTTTTGCACCTACCTTCAAGAAAAACCTTGGATCACGTATTGCT 1381

Db 758 GluHisTyrGluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAla 777
QY 1382 GCTCTAAAGTGATA 1396
Db 778 AlaLeuLysValIle 782

RESULT 4
ABU92026
ID ABU92026 standard; protein; 831 AA.
XX
AC ABU92026;
XX
DT 15-JUL-2003 (first entry)
XX
DE Human protein modification and maintenance molecule-6 (PMMM-6) .
XX
KW Human; protein modification and maintenance molecule; PMMM; cancer;
KW cell proliferation disorder; atherosclerosis; neurological disorder;
KW epilepsy; Huntington's disease; stroke; immune disorder; allergy;
KW inflammatory disorder; AIDS; developmental disorder; hypothyroidism;
KW Cushing's syndrome; gastrointestinal disorder; epithelial disorder;
KW infection; cytostatic; antiarteriosclerotic; anticonvulsant; nootropic;
KW neuroprotective; cerebroprotective; anti-HIV; antiallergic; vulnerary;
KW antiinflammatory; thyromimetic.
XX
OS Homo sapiens.
XX
PN WO2003031939-A2.
XX
PD 17-APR-2003.
XX
PF 11-OCT-2002; 2002WO-US032850.
XX
PR 12-OCT-2001; 2001US-0329689P.
PR 25-OCT-2001; 2001US-0335703P.
PR 09-NOV-2001; 2001US-0348887P.
PR 28-NOV-2001; 2001US-0334145P.
PR 06-DEC-2001; 2001US-0337451P.
PR 14-DEC-2001; 2001US-0340584P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Ramkumar J, Gorvad AE, Baughn MR, Emerling BM, Yang J, Lee SY;
PI Tran UK, Becha SD, Duggan BW, Lee EA, Griffin JA, Li JX;
PI Sprague WW, Hafalia AJA, Chawla NK, Lehr-Mason PM, Kable AE, Yue H;
PI Marquis JP, Yao MG, Richardson TW, Tang TY, Jin P, Chien D;
PI Bhatia U, Burrill JD, Lee S, Blake JJ, Ho A, Zheng W;
XX
DR WPI; 2003-430274/40.
DR N-PSDB; ACA92421.
XX
PT New human protein modification and maintenance molecules (PMMM), useful
PT for diagnosing, treating and preventing diseases or conditions associated
PT with the aberrant PMMM expression e.g. cancer, atherosclerosis, or
PT infections.
XX
PS Claim 1; Page 242-243; 311pp; English.
XX

The present invention relates to the isolation of human protein
CC modification and maintenance molecules (PMMM), and the polynucleotide
CC sequences encoding them. A total of 40 PMMM polypeptides (designated PMMM
CC -1 to PMMM-40) are disclosed. The sequences of the invention are useful
CC for diagnosing a condition or disease associated with the expression of
CC PMMM in a subject, preparing a polyclonal or monoclonal antibody, and
CC generating an expression profile of a sample containing the
CC polynucleotides. The diseases or conditions associated with decreased
CC expression or overexpression of PMMM are cell proliferation disorders
CC (e.g. cancer, atherosclerosis), neurological disorders (e.g. epilepsy,
CC Huntington's disease, stroke), immune/inflammatory disorders, (e.g. AIDS,
CC allergies), developmental disorders (e.g. hypothyroidism, Cushing's
CC syndrome), gastrointestinal or epithelial disorders, and infections. The
CC PMMM polypeptides or their fragments are useful in screening compounds

CC for effectiveness as agonists or antagonists of the polypeptides, or in
CC altering the expression of the target polynucleotide and compounds that
CC specifically bind to, or modulate the activity of the polypeptide.
CC ABU92021-ABU92060 represent the human PMMM polypeptides of the invention
XX
SQ Sequence 831 AA;
Alignment Scores:
Pred. No.: 7,91e-251 Length: 831
Score: 2447.50 Matches: 465
Percent Similarity: 90.5% Conservative: 0
Best Local Similarity: 90.5% Mismatches: 0
Query Match: 82.8% Indels: 49
DB: 6 Gaps: 1

US-10-825-632-6 (1-1669) x ABU92026 (1-831)

QY 2 ACAGGTACACGAATCCTTAAGTCACCTTTTAAAGTGTGACGAATATGATTGATCTGAA 61
Db 318 ThrGlyThrAlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGlu 337
QY 62 GGAAGGATCATAGATGTCATAGATAAGGAACCTAAATCAACCTTTTGAGATTCTATTGAA 121
Db 338 GlyArgIleLeuAspValIleAspLysGluLeuIleGlnProPheGluLeuPheGlu 357
QY 122 GGAGTTCAATATATTGCCAGAGCTGGATGCTCTGAGGAAATATGCTTGTGTCATC 181
Db 358 GlyValGluTyrIleAlaArgAlaGlyTyrProGluGlyLysTyrAlaTyrSerIle 377
QY 182 CTACTAGATCGCTCCAGACTCGCTTACAGATAGTGTGATCTACCTGATTAATTATC 241
Db 378 LeuLeuAspArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIle 397
QY 242 CCAGTAGAAGATGATGTTATGGAAGGAGAGACTCATTGAGTCAGTGCCTGATTCTGTG 301
Db 398 ProValGluAspValMetGluArgGlnArgLeuIleGluSerValProAspSerVal 417
QY 302 ACGCCACTAATTATCTATGAAGAAACAACAGACATCTGGATAAATATCCATGACATCTTT 361
Db 418 ThrProLeuIleIleTyrGluGluThrThrAspIleTyrIleAsnIleHisAspIlePhe 437
QY 362 CATGTTTTTCCCAAGTCACGAAGAGAAATGATTGATTATTTTTCCTCTGAATGCAA 421
Db 438 HisValPheProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLys 457
QY 422 ACAGTTTCGTCATTTATACAAAATTAAGATCTATTATTTAAAGGAAGCAATATAAACA 481
Db 458 ThrGlyPheArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrIleSarg 477
QY 482 TCCAGTGTGGGTGCTGCTCCCAAGTGATTTCAAGTGTCTCTATCAAGAGAGATAGCA 541
Db 478 SerSerGlyGlyLeuProAlaProSerAspPheLysCysProIleLysGluIleAla 497
QY 542 ATTTACAGTGTGAATGGGAAGTTCTTGCCGCGCATGATCTAATATCAAGTTGATGAA 601
Db 498 IleThrSerGlyGluTyrGluValLeuGlyArgHisGlySerAsnIleGlnValAspGlu 517
QY 602 GTCAGAAGGTGGGTATATTTTGAAGGCACCAAGACTCCCTTTTAGACATCACTGTAC 661
Db 518 ValArgArgLeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeuTyr 537
QY 662 GTAGTCAGTTACCGTAAATCTCGAGAGGTGACAAAGCTGACTGACCGGTGCTACAT 721
Db 538 ValValSerTyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHis 557
QY 722 TCTTGTGTCATCAGTCAGACTGCTGACTCTTTTATAGTAAGTATAGTAACCAAGAGAT 781
Db 558 SerCysCysIleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsn 577
QY 782 CCACACTGTGTGTCCTTTTACAGCTATCAAGTCTCTGAAGATGACCCCACTTGCAAAACA 841
Db 578 ProHisCysValSerLeuTyrLysLeuSerSerProGluAspProThrCysLysThr 597

QY 842 AAGAAATTTGGGCCACCAATTTTGATTCAGAGAGTCTCTTCTCTGACTATATCTCTCA 901
 Db LyeGluPheTrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProPro 617
 QY 902 GAAATTTCTCTTTTGAAGTACTACTGGATTTACATTTGATGGATGCTCTCAAGCCT 961
 Db GluIlePheSerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysPro 637
 QY 962 CATGATCTACAGCTCGAAGAAATATCTACTGTGTTGTTTCATATATGTGTCTCTCAG 1021
 Db HisAspLeuGlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGln 657
 QY 1021 ----- 1021
 Db ValGlnLeuValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAla 677
 QY 1021 ----- 1021
 Db SerLeuGlyTyrValValValValLleAspAsnArgGlySerCysHisArgGlyLeuLys 697
 QY 1022 -----GTTGCTATTGCTGGGGCCCGCCAGTCACTCTGTGG 1054
 Db PheGluGlyAlaPheLysTyrLysMetValAlaIleAlaGlyAlaProValThrLeuTrp 717
 QY 1055 ATCTTCTATGATACAGATACAGGAACGTTATATGGTCACTCCGTGACGAGATGACAG 1114
 Db IlePheTyrAspThrGlyTyrThrGluArgTyrMetGlyHisProAspGlnAsnGluGln 737
 QY 1115 GGTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAGTTCCCTCTGAAACCAATCGT 1174
 Db GlyTyrTyrLeuGlySerValAlaMetGlnAlaGluLysPheProSerGluProAsnArg 757
 QY 1175 TTACTGCTCTTACATGGTTTCTCTGGATGAGATGCCATTTTGACATACCATATATTA 1234
 Db LeuLeuLeuLeuHisGlyPheLeuAspGluAsnValHisPheAlaHisThrSerIleLeu 777
 QY 1235 CTGAGTTTTTACTGAGGCTGGAAAGCCATATGATTTTACAGATCTATCTCTCAGGAGAG 1294
 Db LeuSerPheLeuValArgAlaGlyLysProTyrAspLeuGlnIleTyrProGlnGluArg 797
 QY 1295 CACAGCATAAGAGTCTCTCAATCGGAGAACATATTAAGTCACTCTTTTGCACTACCTT 1354
 Db HisSerIleArgValProGluSerGlyGluHisTyrGluLeuHisLeuLeuHisTyrLeu 817
 QY 1355 CAAGAAACCTTGATCAGCTATTGCTGCTCTTAAAGTGATA 1396
 Db 818 GlnGluAsnLeuGlySerArgIleAlaLeuLysValIle 831

RESULT 5
 AAB93565
 ID AAB93565 standard; protein; 632 AA.
 AC AAB93565;
 XX
 XX 26-JUN-2001 (first entry)
 DT Human protein sequence SEQ ID NO:12964.
 DE Human; primer; detection; diagnosis; antisense therapy; gene therapy.
 XX Homo sapiens.
 OS EP1074617-A2.
 FN 07-FEB-2001.
 PD 28-JUL-2000; 2000EP-00116126.
 PF 29-JUL-1999; 99JP-00248036.
 PR 27-AUG-1999; 99JP-00300253.
 PR 11-JAN-2000; 2000JP-00118776.
 PR 02-MAY-2000; 2000JP-00183767.
 PR 09-JUN-2000; 2000JP-00241899.

XX (HELI-) HELIX RES INST.

PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 FI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 DR WPI; 2001-318749/34.

XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.

PS Claim 8; SEQ ID NO 12964; 2537pp + Sequence Listing; English.

XX The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primer sets are useful for synthesizing polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention

XX SQ Sequence 632 AA;

Alignment Scores: 2,97e-250 Length: 632
 Pred. No.: 2441.50 Matches: 464
 Score: 2441.50
 Percent Similarity: 90.3% Conservative: 0
 Best Local Similarity: 90.3% Mismatches: 1
 Query Match: 82.6% Indels: 49
 DB: 4 Gaps: 1

US-10-825-632-6 (1-1669) x AAB93565 (1-632)

QY 2 ACAGGTACAGCAATCTCTAAAGTCACTTTTAAGATGTCAGAAATAATGATTGATCTGAA 61
 Db 119 ThrGlyThrAlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGlu 138
 QY 62 GBAAGGATCATAGATCTCATAGATAGGAACCTTAATTCACCTTTTGGAGATTCTATTGAA 121
 Db 139 GlyArgIleIleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGlu 158
 QY 122 GBAAGGATCATATTCCTCCAGACTGATGAGTCTGAGGAGAAATATGCTTGGTCCATC 181
 Db 159 GlyValGluTyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIle 178
 QY 182 CTACTAGATCGTCCAGACTCGCTCAGATGATGATGATGATGATGATGATGATGATGATGAT 241
 Db 179 LeuLeuAspArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIle 198
 QY 242 CCAAGTAAAGATGATCTTATGAAAGGAGGAGAGACTCATTGATGATGATGATGATGATGATG 301
 Db 199 ProValGluAspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerVal 218
 QY 302 ACGCCACTATTATCTTAAGAAACAAACAGACATCTCGATTAATATCATGATCTTT 361
 Db 219 ThrProLeuIleIleTyrGluThrThrAspIleTrpIleAsnIleHisAspIlePhe 238

QY	362	CATGTTTTCCTCCCAAGTCACGAGAGGAAATTCAGTTTATTTTGTGCTCTGTAATCAAA	1295	CACAGCATAGAGTTCCTGAATCGGGAGAACATATATGAACTGATCTTTTGCACCTACTT	1354
Db	239	HisValPheProGlnSerHisGluGluGluLeuGluPheAlaSerGluCysLys	599	HisSerIleArgValProGluSerGlyGluHisTyrGluLeuHisLeuLeuHisTyrLeu	618
QY	422	ACAGGTTTCCTGTCATTTATACAAAATACATCTATTTAAAGGAAACCAATATTAACGA	1355	CAAGAAACCTTCGATCAGCTATTGCTGCTCTTAAAGTGATA	1396
Db	259	ThrGlyPheArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArg	619	GlnGluAsnLeuGlySerArgIleAlaLeuLysValIle	632
QY	482	TCCAGTGGTGGCTGCTCCTCCAGTGATTTCAAGTGCTCTATCAAGAGGAGATAGCA	RESULT 6		
Db	279	SerSerGlyLeuProIleProSerAspPheLysCysProIleLysGluGluLeuAla	ADT04044		
QY	542	ATTACCAAGCTGTAATGGAAGTCTTTCGCCGCGCATGATCTAATATCCAAGTTGATGAA	ID	ADT04044 standard; protein; 738 AA.	
Db	299	IleThrSerGlyGluPheValLeuGluArgHisGlySerAsnIleGlnValAspGlu	AC	ADT04044;	
QY	602	GTCAAGAGCTGTATATTTGAAGCCACCAAGCTCCCTTTAGAGCATCACCCTGTAC	XX		
Db	319	ValArgArgLeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyr	XX	30-DEC-2004 (first entry)	
QY	662	GTAGTCAGTTACGTAATCCTCGAGAGGTGACAGGCTGACTGACCGTGGCTACTCACAT	XX	Human protein modification and maintenance molecule protein SeqID24.	
Db	339	ValValSerTyrValAsnProGlyGluValThrArgLeuThrAspArgSerTyrSerHis	XX	protein modification and maintenance molecule; PMMM; cytostatic;	
QY	722	TCTTGCTGCATCAGTCAGTCTGACTTCTTTATAGTATAGTAACTAGTAAACCAAGAAAT	XX	immunomodulator; expression; immune disorder; cancer; human.	
Db	359	SerCysCysIleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsn	OS	Homo sapiens.	
QY	782	CCACACTGTGTGTCCTTTACAGCTATCAAGTCTCAGATGACCCCAACTGTGCAAAACA	XX	WO2004084806-A2.	
Db	379	ProHisCysValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThr	XX	07-OCT-2004.	
QY	842	AAGCAATTTGGCCACCAATTTGGATTCAGCAGTCTCTTCTGACTATCTCTCTCCA	PD	16-MAR-2004; 2004WO-US008006.	
Db	399	LysGluPheTrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProPro	PF	21-MAR-2003; 2003US-0456864P.	
QY	902	GAATTTTCTCTTTTCAAGTACTACTGGAATTTACATGTATGGGATGCTCAAGCCT	PR	03-APR-2003; 2003US-0460512P.	
Db	419	GluIlePheSerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysPro	PR	19-MAY-2003; 2003US-0472027P.	
QY	962	CATGATCTACAGCTCGAAGAAATATCTACTGTGCTGTTTCATATATGTTGGTCTCAG	PR	30-MAY-2003; 2003US-0475072P.	
Db	439	HisAspLeuGlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGln	PR	24-JUN-2003; 2003US-0482575P.	
QY	1021	-----	PR	13-AUG-2003; 2003US-0495151P.	
Db	459	ValGlnLeuValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAla	PA	(INCY-) INCYTE CORP.	
QY	1021	-----	XX	Baughn MR, Marquis JP, Kable AE, Chawla NK, Emerling BM, Lee SY;	
Db	479	SerLeuGlyTyrValValValValIleAspAsnArgGlySerCysHisArgGlyLeuLys	PI	Hafalia AJA, Ramkumar J, Richardson TW, Wang JT, Nakamura LL;	
QY	1022	-----GTTGCTATTGTGGGGCCCGACGTCACCTGTGTGG	PI	Yang J, Jin P, Becha SD, Wilson AD	
Db	499	PheGluGlyAlaPheLysTyrLysMetValAlaIleAlaGlyAlaProValThrLeuTrp	XX	WPI; 2004-710256/69.	
QY	1055	ATCTTCTATGATACAGATACACGNACTTATATGGGTTCACCTGACCCAGCAATGACAG	DR	N-PSDB; ADT04072.	
Db	519	IlePheTyrAspThrGlyTyrThrGluArgTyrMetGlyHisProAspGlnAsnGluGln	PT	New human protein modification and maintenance molecules (PMMM)	
QY	1115	GCTATTACTAGTACTGTGGCCATGCAAGCAGAAAAAGTTCCCTCTCGAACCAAAATCGT	PT	polypeptide, useful in preparing a composition for treating a disease	
Db	539	GlyTyrTyrLeuGlySerValAlaMetGlnAlaGluLysPheProSerGluProAsnArg	PT	associated with decreased expression or overexpression of PMMM e.g.,	
QY	1175	TTACTGCTCTTACATGGTTTCTCGATGAGAAATGCCATTTTGCACATACAGTATATTA	XX	cancer.	
Db	559	LeuLeuLeuLeuHisGlyPheLeuAspGluAsnValHisPheAlaHisThrSerIleLeu	PS	Claim 1; SEQ ID NO 24; 222pp; English/	
QY	1235	CTGAGTTTTTGTAGGGCTGGAAGCCATATGATTTTACAGATCTATCTCCAGGAGAGA	CC	This invention relates to novel human protein modification and	
Db	579	LeuSerPheLeuValArgAlaGlyLysProTyrAspLeuGlnIleTyrProGlnGluArg	CC	maintenance molecule (PMMM) proteins and the DNA sequences which encode	
			CC	them. The invention may be useful for the production of compounds with a	
			CC	cytostatic or immunomodulator activity. The proteins are useful in	
			CC	preparing a composition for diagnosing or treating a disease or condition	
			CC	associated with decreased expression or overexpression of functional	
			CC	PMMM, for example immune disorders or cancer. The present sequence is	
			CC	that of a human protein modification and maintenance molecule (PMMM)	
			CC	protein of the invention.	
			SQ	Sequence 738 AA;	
Alignment Scores:					Length: 738
Pred. No.:					Matches: 458
Score:					Conservative: 0
Percent Similarity:					Mismatches: 6
Best Local Similarity:					Indels: 1
Query Match:					Gaps: 1
DB:					8

PT infections.

XX PS Claim 1; Page 249-251; 31pp; English.

XX CC The present invention relates to the isolation of human protein

CC modification and maintenance molecules (PMWM), and the polynucleotide

CC sequences encoding them. A total of 40 PMWM polypeptides (designated PMWM

CC -1 to PMWM-40) are disclosed. The sequences of the invention are useful

CC for diagnosing a condition or disease associated with the expression of

CC PMWM in a subject, preparing a polyclonal or monoclonal antibody, and

CC generating an expression profile of a sample containing the

CC polynucleotides. The diseases or conditions associated with decreased

CC expression or overexpression of PMWM are cell proliferation disorders

CC (e.g. cancer, atherosclerosis), neurological disorders (e.g. epilepsy,

CC Huntington's disease, stroke), immune/inflammatory disorders (e.g. AIDS,

CC allergies), developmental disorders (e.g. hypothyroidism, Cushing's

CC syndrome), gastrointestinal or epithelial disorders, and infections. The

CC PMWM polypeptides or their fragments are useful in screening compounds

CC for effectiveness as agonists or antagonists of the polypeptides, or in

CC altering the expression of the target polynucleotide and compounds that

CC specifically bind to, or modulate the activity of the polypeptide.

CC ABU92021-ABU92060 represent the human PMWM polypeptides of the invention

XX SQ Sequence 824 AA;

Alignment Scores:

Pred. No.:	4,11e-248	Length:	824
Score:	2422.00	Matches:	455
Percent Similarity:	82.3%	Conservative:	0
Best Local Similarity:	82.3%	Mismatches:	0
Query Match:	81.9%	Indels:	100
DB:	6	Gaps:	1

US-10-825-632-6 (1-1669) x ABU92030 (1-824)

Qy 2 ACAGGTACAGCAATCTCTAAAGTCACTTTTAAAGATGTCAGAAATATGATGATCTGAA 61

Db 260 ThrGlyThrAlaAsnProLysValThrPheLysMetSerGluLeuMetIleAspAlaGlu 279

Qy 62 GDAAGGATCATAGATGTCATAGATAAGGAACCTAAATCAACCTTTTGAGATCTCTATTGAA 121

Db 280 GlyArgIleLeuAspValIleAspLysGluLeuIleGlnProPheGluLeuPheGlu 299

Qy 122 GAGATTCAATATATTCAGAGCTGGATGGATCTCTGAGGAAATATGCTTGGTCCATC 181

Db 300 GlyValGluTyrIleAlaArgAlaGlyTyrThrProGluGlyLysTyrAlaTrpSerIle 319

Qy 182 CTACTAGATCGCTCCAGACTCCCTACAGATAGTGTTCATCTCACTGAAATATTATC 241

Db 320 LeuLeuAspArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIle 339

Qy 242 CCAGTAGAAGATGATGTTATGGAAGCGAGACTCATTTGAGTCACTGCTGATTCGTG 301

Db 340 ProValGluAspValMetGluArgGlnArgLeuIleGluSerValProAspSerVal 359

Qy 302 ACGCCACTAATTATCTATGAAGAAACAACAGACATCTGGATAAATATCCATGACATCTTT 361

Db 360 ThrProLeuIleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePhe 379

Qy 362 CNTGTTTTTCCCAAGTCAAGAGAGAAATGAGTTTATTTTTCCTCTCAATGCAAA 421

Db 380 HisValPheProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLys 399

Qy 422 ACAGTTTCGCTCATTTATACAAATATACATCTATTTTAAAGAAAGCAAAATATAACGA 481

Db 400 ThrGlyPheArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArg 419

Qy 482 TCCAGTGGTGGCTGCTCCAGTGATTTTCAAGTGCCTATCAAGAGAGATAGCA 541

Db 420 SerSerGlyGlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAla 439

Qy 542 ATTAACAGTGGTGAATGGGAAGTTCTTGGCCGCATGGATCTAATATCAAGTTGATGAA 601

440 IleThrSerGlyGluTyrGluValLeuGlyArgHisGlySerAsnIleGlnValAspGlu 459

602 GTCAGAGGCTGGTATATTTTGAAGGACCAAGACTCCCTTTAGAGCATCACCTGTAC 661

460 ValArgArgLeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyr 479

662 GTAGTCAGTTACGTAAATCCCTGGAGAGCTGACAAAGGCTGACTGACCGTGGCTACTCACAT 721

480 ValValSerTyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHis 499

722 TCTTGCTGCATCAGTCAGACATCTTCTTTATAGTAAGTATAGTAACACAGAGAAT 781

500 SerCysCysIleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsn 519

782 CCACACTGTGTCTCTTTTACAGCTATCAAGCTCTCTGAAGATGACCCAACTTTGCAAAACA 841

520 ProHisCysValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThr 539

842 AAGGAATTTTGGGCCACCACATTTTGATTCAGCAGGCTCTCTTCTCTGACTATATCTCTCCA 901

540 LysGluPheTrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProPro 559

902 GAAATTTCTCTTTTGAAGTACTGATTCGATTTACATTTGATGGATGCTCTTCAAGGCT 961

560 GluIlePheSerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysPro 579

962 CATGATCTACGCTGGAAAGAAATATCTCTGCTGTCTGTTTCAATATGCTGGTCTCTCAG 1021

580 HisAspLeuGlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGln 599

1021 ----- 1021

600 ValGlnLeuValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAla 619

1021 ----- 1021

620 SerLeuGlyTyrValValValValIleAspAsnArgGlySerCysHisArgGlyLeuLys 639

1021 ----- 1021

640 PheGluGlyAlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGly 659

1021 ----- 1021

660 LeuGlnTyrLeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGly 679

1021 ----- 1021

680 TrpSerTyrGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArg 699

1022 GTTGCTATTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGAA 1081

700 ValAlaIleAlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGlu 719

1082 CGTTATATGGTCCACCTGACCAGAAATGAACAGGCTATTACTTAGGATCTGTGGCCATG 1141

720 ArgTyrMetGlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMet 739

1142 CAAGCAGAAAAGTTCCCTCTCAACCAATCGTTTACTGTCTCTTACATGTTCTCTGGAT 1201

740 GlnAlaGluLysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAsp 759

1202 GAGAAATCCATTTTGCACATACCACTAGTATTTACTAGTGTGTGTGTGTGTGTGTGTGT 1261

760 GluAsnValHisPheAlaHisThrSerIleLeuLeuSerPheLeuValIleAlaGlyLys 779

1262 CCATATGATTTACAGATCTATCTCTCAGGAGACACACATAAGAGTTTCTCGAATCGGGA 1321

780 ProTyrAspLeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGly 799

1322 GAACATTATGAATCGATCTTTTGCACCTACCTTCAAGAAACCTTGGATTCAGCTATTGCT 1381

800 GluHisTyrGluLeuHisLeuHisTyrLeuGluGluAsnLeuGlySerArgIleAla 819

QY	1382	GCTCTAAAGTGATA	1396	Percent Similarity: 82.3%	Conservative: 0
Db	820	AlaLeuLysValIle	824	Best Local Similarity: 82.3%	Mismatches: 0
				Query Match: 81.9%	Indels: 100
				DB:	Gaps: 1
RESULT 8					
AAAB47187				US-10-825-632-6 (1-1669) x AAB47187 (1-882)	
ID	AAAB47187	standard; protein; 882 AA.			
AC	AAAB47187;				
XX					
DT	29-JUN-2001	(first entry)			
XX					
DE	Human DPP8.				
XX					
KW	Human, dipeptidyl aminopeptidase; DPP8; prolly oligopeptidase;				
KW	dipeptidyl peptidase; DPP8; T cell; cleavage; diarrhoea;				
KW	growth hormone deficiency; glucose level; mucosal regeneration;				
KW	non-insulin dependent diabetes mellitus; glucose intolerance;				
XX	immunosuppression.				
OS	Homo sapiens.				
XX					
Key	Location/Qualifiers				
FT	Active-site	739			
FT	/note= "Forms part of Ser-Asp-His catalytic triad"	817			
FT	Active-site	849			
FT	/note= "Forms part of Ser-Asp-His catalytic triad"				
FT	Active-site				
FT	/note= "Forms part of Ser-Asp-His catalytic triad"				
XX					
XX	WO200119866-A1.				
XX					
XX	22-MAR-2001.				
XX					
XX	11-SEP-2000; 2000WO-AU001085.				
XX					
PR	10-SEP-1999; 99AU-00002762.				
PR	18-FEB-2000; 2000AU-00005709.				
XX					
XX	(UNSY) UNIV SYDNEY.				
XX					
XX	Abbott CA, Gorell MD;				
XX	WPI; 2001-281520/29.				
DR	N-PSDB; AAC65694.				
XX					
XX	New human dipeptidyl aminopeptidase (DPP8) useful for cleaving				
PI	substrates, identifying inhibitors of DPP8 catalytic activity which have				
PI	therapeutic uses, and for detecting activated T cells.				
XX					
PS	Claim 1; Fig 2; 78pp; English.				
XX					
XX	This sequence represents human dipeptidyl aminopeptidase (DPP8). DPP8 has				
CC	substrate specificity for H-Ala-Pro-pNA, H-Gly-Pro-pNA and H-Arg-Pro-pNA.				
CC	Therefore, it is a prolly oligopeptidase and a dipeptidyl peptidase,				
CC	because it is capable of hydrolysing the peptide bond C-terminal to Pro				
CC	in each of these compounds. DPP8 is homologous with human DPP8. DPP8 is				
CC	useful for cleaving a substrate, and for detecting an activated T cell				
CC	which involves measuring the level of DPP8 gene expression in a T cell.				
CC	The level of DPP8 expression is detected by detecting the amount of DPP8				
CC	RNA in the cell. It is also useful for identifying a molecule capable of				
CC	inhibiting DPP8 catalytic activity by DPP8. Molecules identified as				
CC	inhibiting DPP8 catalytic activity may be useful for treating diarrhoea,				
CC	growth hormone deficiency, lowering glucose levels in non-insulin				
CC	dependent diabetes mellitus and other disorders involving glucose				
CC	intolerance, enhancing mucosal regeneration and as immunosuppressants				
XX					
SQ	Sequence 882 AA;				
Alignment Scores:					
Pred. No.:	4, 26e-248	Length:	882		
Score:	2422.00	Matches:	465		

QY	2	ACAGGTACAGCAAACTCTAAAGTCACTTTTAAAGATGTCAGAAATAATGATTGATGCTGAA	61
Db	318	ThrGlyThrAlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGlu	337
QY	62	GGAAGGATCATAGATGTCATAGATAAGGAACCTAAATCAACCTTTTGAATTTCTATTGAA	121
Db	338	GlyArgIleIleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGlu	357
QY	122	GGAGTTGAATATATTGCCAGAGCTGGATGGACTCTCTGAGGGGAAAATATGCTTGGTCCATC	181
Db	358	GlyValGluIleAlaArgAlaGlyIleThrProGluGlyLysIleAlaIlePheSerIle	377
QY	182	CTACTAGATCGCTCCAGACTCGCTCAGACATAGTGTGATCTCACCTGAATATTATTC	241
Db	378	LeuLeuAspArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIle	397
QY	242	CCAGTAAAGATGATGTTATGGAAGGACAGAGACTCATTTGAGTCAGTCCCTGATCTGTG	301
Db	398	ProValGluAspValMetGluArgGlnArgLeuIleGluSerValProAspSerVal	417
QY	302	ACGCCACTAATATCTATCAAGAAACACAGACATCTGGATAAATATCCATGACATCTTT	361
Db	418	ThrProLeuIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIle	437
QY	362	CATGTTTTTCCCAAGTCACAGAGAAATGAGTTTATTTTGGCTCTGAAATGAAA	421
Db	438	HisValPheProGlnSerHisGluGluIleGluPheIlePheAlaSerGluCysLys	457
QY	422	ACAGTTTCCGTCATTATACAAATACATCTATTTTAAAGAAACAAATATAACGA	481
Db	458	ThrGlyPheArgHisLeuIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIle	477
QY	482	TCCAGTGGTGGCTGCTCCCTCCAAAGTGATTTCAAGTGCTCTATCAAAGAGGAGATAGCA	541
Db	478	SerSerGlyGlyLeuProAlaProSerAspPheLysCysProIleLysGluIleAla	497
QY	542	ATTACAGTGGTGAATGGAAAGTTCTTGGCCGGATGGATCTATATATCAAGTTGATGAA	601
Db	498	IleThrSerGlyGluIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIle	517
QY	602	GTCAAGAGCTGGTATATTTTGAAGGACCAAGAGCTCCCTTTAGAGCATCACCTGTAC	661
Db	518	ValArgArgLeuValIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIle	537
QY	662	GATGTCAGTTACGTAATCCTGGAGAGGTGACCAAGGCTGACTACCGTGGCTACTCACAT	721
Db	538	ValValSerTyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHis	557
QY	722	TCTTGTGTCATCAGTCAGCAGCTGTGACTTCTTTATAAGTAAGTATAGTACCAAGAAGAT	781
Db	558	SerCysCysIleSerGlnHisCysAspPheIleSerLysIleSerAsnGlnLysAsn	577
QY	782	CCACACTGTGTGCTCCCTTTTCAAGCTATCAAGTCTCAAGTCAAGTCAAGTCAAGTCAAGT	841
Db	578	ProHisCysValSerLeuIleLysLeuSerSerProLysAspProThrCysLysThr	597
QY	842	RAGGATTTTGGGCCACCATTTTGGATTCAGAGTCTCTCTTCTGACTATATCTCTCA	901
Db	598	LysGluPheIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIle	617
QY	902	GAAATTTTCTTTTGAAGTACTACTGGATTTACATTTGATGGATGCTCTCAAGCCT	961
Db	618	GluIlePheSerPheGluSerThrThrGlyPheThrLeuIleIleIleIleIleIleIleIleIleIle	637
QY	962	CATGATCTACAGCTGGAAAGAAATATCTACTGTGCTGTCTATATATATGTTGCTCTCAG	1021

```
Db 638 HisAspLeuGlnProGlyLysTyrProThrValLeuPheIleTyrGlyGlyProGln 657
Qy 1021 ----- 1021
Db 658 ValGlnLeuValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAla 677
Qy 1021 ----- 1021
Db 678 SerLeuGlyTyrValValValValLeuAspAsnArgGlySerCysHisArgGlyLeuLys 697
Qy 1021 ----- 1021
Db 698 PheGluGlyAlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGly 717
Qy 1021 ----- 1021
Db 718 LeuGlnTyrLeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGly 737
Qy 1021 ----- 1021
Db 738 TrpSerTyrGlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArg 757
Qy 1022 GTTGTCTATTGCTGGGGCCCACTCTCTGTGGATCTTCTATGATACAGGATACAGGAA 1081
Db 758 ValAlaIleAlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGlu 777
Qy 1082 CGTTATATGGTCCACCTGACAGCAATGAACAGGGCTATTACTTAGGATCTGTGGCCATG 1141
Db 778 ArgTyrMetGlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMet 797
Qy 1142 CAAGCAGAAAGTCCCTCTCAACCAATCGTTTACTGCTCTTACATGGTTTCTCGGAT 1201
Db 798 GlnAlaGlyLysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAsp 817
Qy 1202 GAGAATGTCATTTGCACATACCAAGTATATTACTGAGTTTTTTAGTGGGCTGGAAG 1261
Db 818 GluAsnValHisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLys 837
Qy 1262 CMTATGATTACAGATCTATCTCAGGAGACACAGCATAAGAGTTCTTGATCGGGA 1321
Db 838 ProTyrAspLeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGly 857
Qy 1322 GACATATCACTGCATCTTTTGCACCTACCTTCAAGAAACCTTGGATCAGCTATTGCT 1381
Db 858 GluHisTyrGluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAla 877
Qy 1382 GCTCTAAAGTGATA 1396
Db 878 AlaLeuLysValIle 882
RESULT 9
ABG61591
ID ABG61591 standard; protein; 882 AA.
AC ABG61591;
XX
XX ABG61591;
XX
DT 12-AUG-2002 (first entry)
XX
DE Human DPPIV related serine protease DPPP-1.
XX
KW Human; serine protease; dipeptidyl peptidase IV-related protein; DPPP;
KW DPPIV; infection; human immunodeficiency virus; HIV-1; HIV-2; pain;
KW diabetes; infertility; obesity; anorexia; Parkinson's disease; stroke;
KW heart failure; hypertension; urinary retention; osteoporosis; cancer;
KW ulcer; allergy; cancer; psychotic disorder; neurological disorder;
KW dyskinesia; reproductive disorder; inflammatory disorder;
KW metabolic disorder.
XX
OS Homo sapiens.
XX
XX WO200231134-A2.
XX
XX 18-APR-2002.
PD
```

```
XX 12-OCT-2001; 2001WO-US031874.
PF
XX
PR 12-OCT-2000; 2000US-0240117P.
XX
PA (FERR ) FERRING BV.
XX
PI Qi S, Akinsanya KO, Riviere PJ, Junien J;
XX
DR WPI; 2002-444178/47.
DR N-PSDB; ABK83322.
XX
PT New dipeptidyl peptidase IV-related proteins and nucleic acids encoding
PT the proteins, useful for treating e.g. fungal, bacterial, protozoan and
PT viral infections, cancers, allergies, neurological disorders, or pain.
XX
PS Claim 17; Fig 1; 113pp; English.
XX
CC The present invention relates to the isolation of novel human serine
CC proteases referred to as dipeptidyl peptidase IV (DPPIV)-related proteins
CC (DPPP). The dipeptidyl peptidase IV-related proteins (DPPP) and nucleic
CC acids encoding them are useful for treating infections such as fungal,
CC bacterial, protozoan and viral infections, particularly infections caused
CC by human immunodeficiency virus (HIV-1 or HIV-2), pain, diabetes,
CC precocious puberty, infertility, obesity, anorexia, bulimia, Parkinson's
CC disease, acute heart failure, hypotension, hypertension, urinary
CC retention, osteoporosis, angina pectoris, stroke, ulcers, asthma,
CC allergies, cancers, migraine, vomiting, psychotic and neurological
CC disorders (e.g. anxiety, dementia, or schizophrenia), and dyskinesias.
CC These may also be used in discovering therapeutic agents for the
CC treatment of reproductive, inflammatory and metabolic disorders. ABG61591
CC -ABG61612 represent human DPPP proteins
XX
SQ Sequence 882 AA;
Alignment Scores:
Pred. No.: 4,26e-248 Length: 882
Score: 2422.00 Matches: 465
Percent Similarity: 82.3% Conservative: 0
Best Local Similarity: 82.3% Mismatches: 0
Query Match: 81.9% Indels: 100
DB: 5 Gaps: 1
US-10-825-632-6 (1-1669) x ABG61591 (1-882)
Qy 2 ACAGGTACACCAATCTCTAAAGTCACCTTTAAGATGTCAGAAATTAATGATGCTGAA 61
Db 318 ThrGlyThrAlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGlu 337
Qy 62 GGAAGGATCATAGATGCTCATAGATAGGAACTAATTCACCTTTTGAGATCTATTGAA 121
Db 338 GlyArgIleIleAspValIleAspLysGluLeuIleGlnProPheGluIlePheGlu 357
Qy 122 GGAGTTCAATATATTCAGAGCTGGATGCTCTCTGAGGAGAAATATGCTTGTGTCATC 181
Db 358 GlyValGluTyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIle 377
Qy 182 CTACTAGATCGCTCCAGACTCGCCCTACAGATAGTGTGATCTCACCTGAAATTTATTC 241
Db 378 LeuLeuAspArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIle 397
Qy 242 CCAGTAGAAGATGATGTTATGGAAGGACAGACTCATGAGTCAGTCCCTGATTCGTGTG 301
Db 398 ProValGluAspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerVal 417
Qy 302 ACGCCACTAATATCTATGAGAGAAACAAACAGACATCTGGATTAATATCATGACATCTTT 361
Db 418 ThrProLeuIleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePhe 437
Qy 362 CATGTTTTTCCCAAGTACAGAGAGAAATGAGTTTATTTTTCCTCTGATGCAAA 421
Db 438 HisValPheProGlnSerHisGluGluGluIleGluPheIleAlaSerGluCysLys 457
```

QY 422 ACAGGTTTCGGTCAATTTATACAAATTTACATCTATTTTAAAGGAAGCAATATATAACGA 481
Db 458 ThrGlyPheArgHisLeuTyrIleThrSerIleLeuLysGluSerLysTyrLysArg 477
QY 482 TCCAGTGGTGGCTGCTGCTCCAAAGTATTTCAAGTGTCTATCAAGAGGAGATAGCA 541
Db 478 SerSerGlyGlyLeuProAlaProSerAspPheLysCysProIleLysGluIleAla 497
QY 542 ATTACCAAGTGGTCAATGGGAAGTCTTGGCCGCGCATGATCTCAATATATCAAGTGTATGAA 601
Db 498 IleThrSerGlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGlu 517
QY 602 CTCAGAGGCTGTATATTTTGAAGCCACCAAGACTCCCTTTAGAGCATCACTGTAC 661
Db 518 ValArgGluValTyrPheGluGlyThrLysAspSerProLeuGluHisLeuTyr 537
QY 662 GTAGTCAGTACGTAATCTCGAGAGGTGACAGGCTGACTGACCGTGGCTACTCACAT 721
Db 538 ValValSerTyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHis 557
QY 722 TCTGTGTCATCAGTCAGCACTGTGACTTCTTTATAGTAAAGTAAAGTAAACCAAGAAT 781
Db 558 SerCysCysIleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsn 577
QY 782 CCACACTGTGTGCTCCCTTTACAGCTATCAAGTCTCTGAAGATGACCCAACTTGCAAAACA 841
Db 578 ProHisCysValSerLeuTyrLysLeuSerSerProGluAspProThrCysLysThr 597
QY 842 AAGGAATTTGGGCCACCACTTTGGATTGAGTTCAGCAGGTCTCTTCTGACTATATCTCTCA 901
Db 598 LysGluPheTrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProPro 617
QY 902 GAAATTTCTCTTTTGAAGTACTACTGGATTTCATTTGATGGATGCTCTACAGCCT 961
Db 618 GluIlePheSerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysPro 637
QY 962 CATGATCTACGCTGGAAGAAATCTCTACTGCTGTGTTATATATATGCTGCTCTAG 1021
Db 638 HisAspLeuGlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGln 657
QY 1021 ----- 1021
Db 658 ValGlnLeuValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAla 677
QY 1021 ----- 1021
Db 678 SerLeuGlyTyrValValValIleAspAsnArgGlySerCysHisArgGlyLeuLys 697
QY 1021 ----- 1021
Db 698 PheGluGlyAlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGly 717
QY 1021 ----- 1021
Db 718 LeuGlnTyrLeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGly 737
QY 1021 ----- 1021
Db 738 TrpSerTyrGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArg 757
QY 1022 GTTGCTATTGCTGGGCCCCAGTCACTCTGGGATCTTCTATGATACAGATACACGGA 1081
Db 758 ValAlaIleAlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGlu 777
QY 1082 CGTTATATGGTCCACCTGACCAAGTACAGGGCTATTACTAGGATCTGGGCCATG 1141
Db 778 ArgTyrMetGlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMet 797
QY 1142 CAAGCAGAAAGTTCCTCTGACCAAACTGTTACTGCTTACTGCTTACTGTTCTGGAT 1201
Db 798 GlnAlaGluLysPheProSerGluProAsnArgLeuLeuHisGlyPheLeuAsp 817
QY 1202 GAGAATGTCCATTTTGCACATACCAAGTATATTACTAGTGTCTTTAGTGAGGCGCTGGAAG 1261

Db 818 GluAsnValHisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLys 837
QY 1262 CCATATGATTTACAGATCTATCTCTAGAGAGACACACATAGAGATTCCTGAATCGGA 1321
Db 838 ProTyrAspLeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGly 857
QY 1322 GAACATTATGAACATGATCTTTTGCCTACTCTTCAAGAAAACCTTGGATCACTATTGCT 1381
Db 858 GluHisTyrGlnLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAla 877
QY 1382 GCTCTAAAAGTGATA 1396
Db 878 AlaLeuLysValIle 882
RESULT 10
ID AAG78415 standard; protein; 882 AA.
XX AAG78415;
AC AAG78415;
DT 12-APR-2002 (first entry)
XX Amino acid sequence of 21953 human prollyl oligopeptidase.
DE 21953 prollyl oligopeptidase; antibody; proline; endopeptidase; cancer;
KW cardiovascular disease; autoimmune disease; atopic allergy;
KW neuronal disorder; vascular disorder; prostate disorder; cytostatic;
KW antidiabetic; antiarthritic; antiasthmatic; antiinflammatory;
KW diabetes mellitus; arthritis; multiple sclerosis; asthma;
KW Grave's disease; neuronal disorder; demyelinating disease.
XX Homo sapiens.
XX WO200179473-A2.
XX 25-OCT-2001.
XX 11-APR-2001; 2001WO-US040483.
PF 11-APR-2001; 2000US-0197508P.
XX 18-APR-2001; 2000US-0197508P.
XX (MILL-) MILLENNIUM PHARM INC.
FA Meyers RA, Williamson M;
PI WPI; 2002-034353/04.
XX N-PSDB; AAH99934.
DR New polypeptides 21953, member of human prollyl oligopeptidase family,
PT useful as diagnostic targets and therapeutic agents for controlling
PT cancer, lymphoma and leukemia.
XX Claim 1; Page 102-103; 121pp; English.
XX This invention relates to an isolated 21953 human prollyl oligopeptidase.
CC Which is cytostatic, antidiabetic, antiarthritic, neuroprotective,
CC antichryoid, dermatological, antipsoriatic, antiasthmatic,
CC ophthalmological, antiinflammatory, vasotropic, antianginal, cardian,
CC anticonvulsant, gynaecological, anorectic and metabolic in its action. Uses include
CC antithyrosclerotic anorectic and metabolic in its action. Uses include
CC gene therapy, expression or activity of 21953 protein modulator, it is
CC useful for identifying a compound which binds to it and can be used in
CC preventing treating or detecting a cellular proliferative or
CC differentiative disorder. The 21953 molecules can act as novel diagnostic
CC targets and therapeutic agents for controlling disorders associated with
CC the aberrant activity or degradation of peptide hormones e.g., disorders
CC associated with cell differentiation, neurological and proliferation such as cancer,
CC immune function, reproductive, neurological and cardiovascular function.
CC The 21953 molecules are thus useful for treating and preventing cellular
CC proliferative and differentiative disorders, haematopoietic neoplastic
CC disorders, immune disorders such as autoimmune diseases, diabetes
CC mellitus, arthritis, multiple sclerosis, asthma, Grave's disease,

CC	neural disorders, demyelinating diseases, vascular disorders and
CC	metabolism or pain disorders. This sequence represents the amino acid
CC	sequence of 21953 human prollyl oligopeptidase
XX	
SQ	Sequence 882 AA;
Alignment Scores:	
Pred. No.:	4,26e-248 Length: 882
Score:	2422.00 Matches: 465
Percent Similarity:	82.3% Conservative: 0
Best Local Similarity:	82.3% Mismatches: 0
Query Match:	81.9% Indels: 100
DB:	5 Gaps: 1
US-10-825-632-6 (1-1669) x AAG78415 (1-882)	
QY	2 AGAGTACAGCAATCTTAAGTCACTTTAAGATGTCAGAAATATGATGCTGAA 61
DB	318 ThrGlyThrAlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGlu 337
QY	62 GGAAGATCATAGATGTCATAGATAAGGAATCACTTTTCAAGCTTTTGGATTTTATTTGAA 121
DB	338 GlyArgIleAspValIleAspLysGluIleGlnProPheGluIleLeuPheGlu 357
QY	122 GGAAGTGAATATATTGCCAGAGCTGGATGGAATCTCTGAGGGAATATATGCTTGGTCCATC 181
DB	358 GlyValGluTyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIle 377
QY	182 CTACTAGATCGCTCCAGACTCGCTCAGATAGTGTGATCTCACTGATTAATTTATTC 241
DB	378 LeuLeuAspArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIle 397
QY	242 CAGTGAAGATGATGTTATGGAAGGAGAGACTCATTGAGTCAGTCCCTGATTTCTG 301
DB	398 ProValGluAspValMetGluArgGlnArgLeuIleGluSerValProAspSerVal 417
QY	302 ACGCCACTAATATCTATGAAGAAACACAGACTCTGGATAAATATCCATGATCATCTTT 361
DB	418 ThrProLeuIleIleTyrGluIuThrThrAspIleTrpIleAsnIleHisAspIlePhe 437
QY	362 CATGTTTTTCCCAAGTCACAGAGGAAATTCAGTTTATTTTGTCTCTGAATCAAA 421
DB	438 HisValPheProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLys 457
QY	422 ACAGGTTCCGTCATTATATCAAAATTCATCTATTTTAAAGGAAACAAATATAACGA 481
DB	458 ThrGlyPheArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArg 477
QY	482 TCAGTGGTGGGCTGCTCTCCAGTGAATTCAGTGTCTATCAAGAGGAGATAGCA 541
DB	478 SerSerGlyGlyLeuProAlaProSerAspPheLysCysProIleLysGluIleAla 497
QY	542 ATTACAGTGGTGAATGGGAAGTTCTTGGCCGATGATCTAATATCCAAAGTTGATGAA 601
DB	498 IleThrSerGlyIuTrpGluValLeuGluGlyArgHisGlySerAsnIleGlnValAspGlu 517
QY	602 GTCAAGGCTGTATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACTGTAC 661
DB	518 ValArgArgLeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeuTyr 537
QY	662 GTAGTCAGTTACGTAATCTCGAGAGGTGACAAAGCTGACTGACCGTGGCTACTCAT 721
DB	538 ValValSerTyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHis 557
QY	722 TCTGTGTCATAGTCAGACATGCTGACTTCTTTTATAGTAAATAGTATAGTACCAGAAAT 781
DB	558 SerCysCysIleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsn 577
QY	782 CCACACTGTGTCCCTTACAGCTATCAAGTCTGAGATGACCCCACTTGCAAAACA 841
DB	578 ProHisCysValSerLeuTyrLysLeuSerSerProGluIuAspAspProThrCysLysThr 597
QY	842 AAGGAATTTGGGCCCACTTTTGGATTTCAGCAGGTCTCTTCTGACTATATCTCTCCA 901

DB	598 LysGluPheTrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProPro 617
QY	902 GAAATTTCTCTTTTAAAGTACTACTGGATTTACATTTGATGGATGCTCTCAAGCCT 961
DB	618 GluIlePheSerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysPro 637
QY	962 CATGATCTACAGCTCGGAAAGAAATATCTACTCTGCTGTGTATATATATGCTGCTCAG 1021
DB	638 HisAspLeuGlnProGlyLysTyrProThrValLeuPheIleTyrGlyGlyProGln 657
QY	1021 ----- 1021
DB	658 ValGlnLeuValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAla 677
QY	1021 ----- 1021
DB	678 SerLeuGlyTyrValValValIleAspAsnArgGlySerCysHisArgGlyLeuLys 697
QY	1021 ----- 1021
DB	698 PheGluGlyAlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGly 717
QY	1021 ----- 1021
DB	718 LeuGlnTyrLeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGly 737
QY	1021 ----- 1021
DB	738 TrpSerTyrGlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArg 757
QY	1022 GTTGCTATTGCTGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGAA 1081
DB	758 ValAlaIleAlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGlu 777
QY	1082 CGTTATATGGTCCCTGACAGAAATGACAGGGCTATTACTTAGGATCTGTGGCCATG 1141
DB	778 ArgTyrMetGlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMet 797
QY	1142 CAGCGAAGAGTTCCCTCTGAACCAATCGTTTACTGCTCTTACATGGTTTCTCGAT 1201
DB	798 GluAlaGluLysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAsp 817
QY	1202 GAGAATGCTCCATTTCCACATACCACTATATTACTAGTGTATTTTAGTGAGGCTGGAAG 1261
DB	818 GluAsnValHisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLys 837
QY	1262 CCATATGATTTACAGATCTATCTCTCAGGAGACACACATAGAGTTCTCGAATCGGGA 1321
DB	838 ProTyrAspLeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGly 857
QY	1322 GAACATTATGAATGATCTTTTGCATCTTCTTCAAGAAACCTTGGATCAGTATTGCT 1381
DB	858 GluHisTyrGluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAla 877
QY	1382 GCTCTAAAGTGATA 1396
DB	878 AlaLeuLysValIle 882
RESULT 11	
AAE24170	standard; protein; 882 AA.
ID	AAE24170
XX	AAE24170;
AC	AAE24170;
XX	23-SEP-2002 (first entry)
DT	Human dipeptidyl peptidase 8 (DPP8) protein.
DE	
XX	Human; dipeptidyl peptidase; DPP; neoplasia; type II diabetes; cirrhosis;
KW	autoimmunity; human immuno deficiency virus; HIV infection; cytostatic;
KW	graft rejection; antidiabetic; antiinflammatory; immunosuppressive;
KW	antiviral; enzyme.

```

XX OS Homo sapiens.
XX FN WO200234900-A1.
XX PD 02-MAY-2002.
XX PF 29-OCT-2001; 2001WO-AU001388.
XX PR 27-OCT-2000; 2000AU-00001078.
XX PA (UNSY ) UNIV SYDNEY.
XX PI Abbott CA, Gorrell MD;
XX DR WPI: 2002-454646/48.
XX DR N-PSDB; AAD38956.
XX CC
XX CC The present invention relates to dipeptidyl peptidase (DPP) proteins and
XX CC polynucleotides encoding such proteins. The DPP peptides are useful for
XX CC screening inhibitors of DPP catalytic activity. The inhibitors are useful
XX CC for treating neoplasia, type II diabetes, cirrhosis, autoimmunity, graft
XX CC rejection and HIV (human immuno deficiency virus) infection. The present
XX CC sequence is human DPP8 protein
XX SQ Sequence 882 AA;

Alignment Scores:
Pred. No.: 4,26e-248 Length: 882
Score: 2422.00 Matches: 465
Percent Similarity: 82.3% Conservative: 0
Best Local Similarity: 82.3% Mismatches: 0
Query Match: 81.9% Indels: 100
DB: 5 Gaps: 1

US-10-825-632-6 (1-1669) x AAE24170 (1-882)
QY 2 ACAGGTACAGCMAATCCTAAAGTCATTTTAAGATGTCAGAAATAATGATGCTGAA 61
Db 318 ThrGlyThrAlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGlu 337
QY 62 GGAAGGATCATAGATGTCATAGATAAGGAATAATCAACCTTTTGAGATTCATTGAA 121
Db 338 GlyArgIleIleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGlu 357
QY 122 GGAGTTGAATATATTGGCCAGAGCTGGATGGACTCCTGAGGAAATAATGCTTGGTCCATC 181
Db 358 GlyValGluIleIleAlaAlaArgAlaGlyTrpThrProGluGlyLysTrpAlaTrpSerIle 377
QY 182 CTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAATTTATTC 241
Db 378 LeuLeuAspArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIle 397
QY 242 CCAGTAGAAGATGATGTTATGGAAGGACAGACTCATGAGTCAGTCCCTGATTCGTG 301
Db 398 ProValGluAspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerVal 417
QY 302 ACGCCTCAATATTATGATGAAGAAACAACAGACATCTGGATAAATATCATGACATCTTT 361
Db 418 ThrProLeuIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIle 437
QY 362 CATCTTTTCCCAAGTACAGAGGAAATGAGTTATTTTGGCTCTGCAATGCAAA 421
Db 438 HisValPheProGlnSerHisGluGluIleGluPheIlePheAlaSerGluCysLys 457
QY 422 ACAGGTTCCGTCAATTTATACAAATATACATCTATTTTAAAGGAAAGCAATATAACGA 481

```

```

Db 458 ThrGlyPheArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArg 477
QY 482 TCCAGTGGTGGCTGCTCTCAAGTGATTTCAAGTGTCTCTATCAAGAGAGAGATAGCA 541
Db 478 SerSerGlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAla 497
QY 542 ATTACCAAGTGGTGAATGGGAAGTCTTGGCCGGCATGGATCTAATATATCAAGTTGATCAA 601
Db 498 IleThrSerGlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGlu 517
QY 602 GTCAGAAGGCTGTATATTTGAAGGCACCAAGACTCCCTTTTAGACATCACTGTGAC 661
Db 518 ValArgArgLeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeuTyr 537
QY 662 GTAGTCAGTTACGTAAATCTCGAGAGGTGACAAGCTGACTGACCGGGCTACTCCAT 721
Db 538 ValValSerTyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHis 557
QY 722 TCTTGCTGATCAGTCAGCACGTGCTCTTATTAAGTAAGTATAGTAAACCAAGAAAT 781
Db 558 SerCysCysIleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsn 577
QY 782 CCACACTGTGTCTCTTACAGCTATCAAGCTATCAAGTCTCGAAGATGACCCAACTGCAAAACA 841
Db 578 ProHisCysValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThr 597
QY 842 AAGGAATTTGGGCCACCACTTTTGATTCACAGAGTCTCTTCTGACTATATCTCTCCA 901
Db 598 LysGluPheTrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProPro 617
QY 902 GAAATTTCTCTTTGAAAGTACTACTGATTTTACATTGTATGGGATGCTCTCAAGCCT 961
Db 618 GluIlePheSerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysPro 637
QY 962 CATGATCTACAGCTCGAAAGAAATATCTACTGCTGTGTTCATATATGCTGTCTCTCAG 1021
Db 638 HisAspLeuGlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGln 657
QY 1021 ----- 1021
Db 658 ValGlnLeuValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAla 677
QY 1021 ----- 1021
Db 678 SerLeuGlyTyrValValValValIleAspAsnArgGlySerCysHisArgGlyLeuLys 697
QY 1021 ----- 1021
Db 698 PheGluGlyAlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGly 717
QY 1021 ----- 1021
Db 718 LeuGlnTyrLeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGly 737
QY 1021 ----- 1021
Db 738 TrpSerTyrGlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArg 757
QY 1022 GTTGCTATTGCTGGGGCCCAAGTCATCTCTGTGGATCTTCTATGATACAGATACACGGAA 1081
Db 758 ValAlaIleAlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGlu 777
QY 1082 GCTTATATGGTCCACCTGACCAAGATGAACAGGGCTATTACTTAGGATCTCTGGCCCATG 1141
Db 778 ArgTyrMetGlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMet 797
QY 1142 CAAGCAGAAAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCTTACATCGTTTCTGTTGAT 1201
Db 798 GlnAlaGluLysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAsp 817
QY 1202 GAGATGTCCATTTTGGCACATACAGTATATTACTAGTTTTTGTAGTGGGCTCGAAG 1261

```


Db 818 GluAsnValHisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLys 837
 QY 1262 CCAATATGATTTACAGATCTATCTCCAGGAGAGACACAGCATAAGAGTCTCTGAATCGGGA 1321
 Db 838 ProTyrAspLeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGly 857
 QY 1322 GAACATTATGAATCCTCATCTTTTGCATCTTACCTTCAAGAAAACCTTGGATCAGTATTGCT 1381
 Db 858 GluHisTyrGluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAla 877
 QY 1382 GCTCTAAAGTCATA 1396
 Db 878 AlaLeuLysValIle 882

RESULT 12
 AAU74749
 ID AAU74749 standard; protein; 882 AA.
 XX
 AC AAU74749;
 XX
 DT 09-APR-2002 (first entry)
 XX
 DE Human protease PRTS-9 protein sequence.
 XX
 KW Human; protease; PRTS; gastrointestinal; Crohn's disease; cancer;
 KW cardiovascular; atherosclerosis; autoimmune disorder; dermatitis;
 KW inflammatory disorder; acquired immunodeficiency syndrome; AIDS;
 KW cell proliferative disorder; developmental disorder; epilepsy;
 KW Duchenne muscular dystrophy; epithelial disorder; neurological disorder;
 KW reproductive disorder; endometriosis.
 XX
 OS Homo sapiens.
 XX
 PN WO200198468-A2.
 XX
 PD 27-DEC-2001.
 XX
 PF 13-JUN-2001; 2001WO-US019178.
 XX
 PR 16-JUN-2000; 2000US-0212336P.
 PR 22-JUN-2000; 2000US-0213955P.
 PR 29-JUN-2000; 2000US-0215396P.
 PR 07-JUL-2000; 2000US-0216821P.
 PR 14-JUL-2000; 2000US-0218946P.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 PI Yue H, Elliott VS, Gandhi AR, Lal P, Au-Young J, Tribouley CM;
 PI Delegeane AM, Baughn MR, Nguyen DB, Lee EA, Hafalia A, Khan FA;
 PI Walia NK, Yao MG, Lu DAM, Patterson C, Tang YT, Walsh RT;
 PI Azimzai Y, Lu Y, Ramkumar J, Xu Y, Reddy R, Das D, Kearney L;
 PI Kallick DA;
 XX
 WP: 2002-090437/12.
 DR N-PSDB; ABK12892.
 XX
 PT Twenty one human proteases (referred to as PRTS-1 to PRTS-21), useful in
 PT the diagnosis, treatment and prevention of gastrointestinal (e.g.
 PT gastritis), cardiovascular (e.g. atherosclerosis) and cell proliferative
 PT (e.g. cancer) disorders.
 XX
 PS Claim 1; Page 140-142; 177pp; English.
 XX
 CC The present invention relates to twenty one new human proteases, referred
 CC to as PRTS-1 to PRTS-21. The PRTS polynucleotides and polypeptides of the
 CC invention are useful in the diagnosis, treatment and prevention of
 CC gastrointestinal e.g. gastritis, esophageal carcinoma and Crohn's
 CC disease, cardiovascular e.g. atherosclerosis, hypertension and myocardial
 CC infarction; autoimmune/inflammatory e.g. acquired immunodeficiency
 CC syndrome (AIDS), allergies and osteoarthritis, cell proliferative e.g.
 CC cancer, developmental e.g. Duchenne and Becker muscular dystrophy,
 CC epithelial e.g. dermatitis, neurological e.g. epilepsy and Alzheimer's
 CC disease and reproductive e.g. infertility and endometriosis disorders.

CC Numerous other examples of each disorder are given in the specification.
 CC The present protein sequence represents the human protease PRTS-9 protein
 CC of the invention
 XX
 SQ Sequence 882 AA;
 Alignment Scores:
 Pred. No.: 4,266-248 Length: 882
 Score: 2422.00 Matches: 465
 Percent Similarity: 82.3% Conservative: 0
 Best Local Similarity: 82.3% Mismatches: 0
 Query Match: 81.9% Indels: 100
 DB: 5 Gaps: 1
 US-10-825-632-6 (1-1669) x AAU74749 (1-882)
 QY 2 ACAGGTACAGCAAACTCTAAAGTCACTTTTAAGATGTGAGAAATAATGATGATGCTGAA 61
 Db 318 ThrGlyThrAlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGlu 337
 QY 62 GGAAGGATCATAGATGTCATAGATAAGAACTAATTAACCTTTTGAGATTCATTGAA 121
 Db 338 GlyArgIleIleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGlu 357
 QY 122 GGAAGTGAATATATGCCAGAGCTGGATGGACTCTGAGGGGAAAATATGCTTGCTCATC 181
 Db 358 GlyValGluTyrIleAlaArgAlaGlyThrProGluGlyLysTyrAlaIrrSerIle 377
 QY 182 CTACTAGATCGCTCCAGACTCGCTCAGATAGTGTGATCTCACCTGAAATATTATTC 241
 Db 378 LeuLeuAspArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIle 397
 QY 242 CCAGTAGAAGATGATGTTATGAAAGGAGAGAGACTCATGAGTGTGAGTGCCTGATCTGTG 301
 Db 398 ProValGluAspValMetGluArgGlnArgLeuIleGluSerValProAspSerVal 417
 QY 302 ACCGCACTAATATCTATGAAGAAACAACAGACATCTCGATAAATATCATGACATCTTT 361
 Db 418 ThrProLeuIleIleTyrGluGluThrAspIleIrrPheAsnIleHisAspIlePhe 437
 QY 362 CATGTTTTTCCCAAGTCAAGAGAGAAATGAGTTTATTTTTCCTCTGAAATGCAAA 421
 Db 438 HisValPheProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLys 457
 QY 422 ACAGGTTCCGTCATTTATACAAAATTAATCTATTTTAAAGGAAAGCAATATAACGA 481
 Db 458 ThrGlyPheArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArg 477
 QY 482 TCCAGTGGTGGCTGCTCTCCAGTCAATTTCAAGTGTCTCTATCAAGAGAGAGATAGCA 541
 Db 478 SerSerGlyGlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAla 497
 QY 542 ATTACAGTGGTGAATGGGAAGTTCTTGCCCGCATGGATCTAATATATCAAGTTGATGAA 601
 Db 498 IleThrSerGlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGlu 517
 QY 602 GTCAGAGGCTGTATATTTTGAAGGCAACAAAGACTCCCTTTTAGACATCACTGTAC 661
 Db 518 ValArgArgLeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeuTyr 537
 QY 662 GTAGTCAGTTACGTAATCTCGAGAGGGTGCACAGGCTGACTGACCGTGGCTACTCAT 721
 Db 538 ValValSerTyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHis 557
 QY 722 TCTTGCTGCATCAGTCAGCAGCTGTGACTCTTTTATAAGTAAAGTAAAGTAAAGTAAAGT 781
 Db 558 SerCysCysIleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsn 577
 QY 782 CCACACTGTGTGCTCCCTTTACAGCTATCAAGCTATCAAGTCTGAGATGACCCCACTTGCAACA 841
 Db 578 ProHisCysValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThr 597
 QY 842 AAGAAATTTGGGCCACCATTTTGGATTTCAGAGGTCCTCTCTCCGTACTATATCTCTCCA 901

Db 598 LysGluPheTrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProPro 617
 Qy 902 GAAATTTCTCTTTTCAAGTACTACTGGATTACATGATGGATGCTCTACAGCCT 961
 Db 618 GluIlePheSerPheGluSerThrThrGlyPheThrLeuIleGlyMetLeuIleYrLysPro 637
 Qy 962 CATGATCTACAGCTCGGAAGAAATATCTCTACTGTCTGTTCATATATGTTGCTCTCAG 1021
 Db 638 HisAspLeuGlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGln 657
 Qy 1021 ----- 1021
 Db 658 ValGlnLeuValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAla 677
 Qy 1021 ----- 1021
 Db 678 SerLeuGlyTyrValValValValIleAspAsnArgGlySerCysHisArgGlyLeuLys 697
 Qy 1021 ----- 1021
 Db 698 PheGluGlyAlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGly 717
 Qy 1021 ----- 1021
 Db 718 LeuGlnTyrLeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGly 737
 Qy 1021 ----- 1021
 Db 738 TrpSerTyrGlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArg 757
 Qy 1022 GTTGCTATTGCTGGGCCCCAGTCATCTGTGGATCTCTATCATATACAGGATACAGGAA 1081
 Db 758 ValAlaIleAlaGlyAlaProValThrLeuIlePheIlePheIlePheIlePheIlePheIle 777
 Qy 1082 CATTATAGGTCACCTCACAGATGACAGATGACAGGCTATTACTAGGATCTGTGGCCATG 1141
 Db 778 ArgTyrMetGlyHisPheAspGlnAsnGlnGlyTyrTyrLeuGlySerValAlaMet 797
 Qy 1142 CAACAGAAAGTCCCTCTGAAACCAATGTTACTGCTTACATGTTCTCTGAT 1201
 Db 798 GlnAlaGluLysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAsp 817
 Qy 1202 GAGAAATGCCATTTTGCACATACACAGTATATTACTGAGTGTGAGGCTGGAAG 1261
 Db 818 GluAsnValHisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLys 837
 Qy 1262 CCATATGATTTACAGATCTATCTCAGGAGACACAGCATAAGATTCCTCAATCGGA 1321
 Db 838 ProTyrAspLeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGly 857
 Qy 1322 GAACATTATGAATGCTCTTTTGCATCTTACCTCAAGAAACCTTGATCAGTATGCT 1381
 Db 858 GluHisTyrGluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAla 877
 Qy 1382 GCTCTAAAGTGATA 1396
 Db 878 AlaLeuLysValIle 882
 RESULT 13
 ID AD117086
 XX AD117086 standard; protein; 882 AA.
 AC
 XX AD117086;
 DT 15-APR-2004 (first entry)
 XX
 DE Human NOVX protein homologue SeqID 622.
 XX
 KW human; NOVX; cardiomyopathy; atherosclerosis; cancer; diabetes;
 KW inflammation; autoimmune disorder; allergy; blood disorder;
 KW acquired immunodeficiency syndrome; AIDS; obesity; asthma;
 KW immunoglobulin (Ig)A nephropathy; cirrhosis; arthritis;

KW Alzheimer's disease; infection; str.
 XX Homo sapiens.
 OS WO200268649-A2.
 PN
 XX
 PD 06-SEP-2002.
 PF
 XX
 PP 31-JAN-2002; 2002WO-US002785.
 PR 31-JAN-2001; 2001US-0265395P.
 PR 31-JAN-2001; 2001US-0265412P.
 PR 31-JAN-2001; 2001US-0265514P.
 PR 31-JAN-2001; 2001US-0265517P.
 PR 02-FEB-2001; 2001US-0266406P.
 PR 05-FEB-2001; 2001US-0266767P.
 PR 07-FEB-2001; 2001US-0266975P.
 PR 07-FEB-2001; 2001US-0267057P.
 PR 08-FEB-2001; 2001US-0267459P.
 PR 09-FEB-2001; 2001US-0267823P.
 PR 15-FEB-2001; 2001US-0268974P.
 PR 26-FEB-2001; 2001US-0271664P.
 PR 27-FEB-2001; 2001US-0271839P.
 PR 27-FEB-2001; 2001US-0271855P.
 PR 02-MAR-2001; 2001US-0272788P.
 PR 02-MAR-2001; 2001US-0273046P.
 PR 14-MAR-2001; 2001US-0275925P.
 PR 14-MAR-2001; 2001US-0275947P.
 PR 14-MAR-2001; 2001US-0275950P.
 PR 15-MAR-2001; 2001US-0275989P.
 PR 15-MAR-2001; 2001US-0276448P.
 PR 16-MAR-2001; 2001US-0276450P.
 PR 16-MAR-2001; 2001US-0276397P.
 PR 16-MAR-2001; 2001US-0276768P.
 PR 20-MAR-2001; 2001US-0278652P.
 PR 26-MAR-2001; 2001US-0278757P.
 PR 26-MAR-2001; 2001US-0278778P.
 PR 29-MAR-2001; 2001US-0279882P.
 PR 30-MAR-2001; 2001US-0279884P.
 PR 30-MAR-2001; 2001US-0280147P.
 PR 11-APR-2001; 2001US-0282992P.
 PR 11-APR-2001; 2001US-0283083P.
 PR 20-APR-2001; 2001US-0285133P.
 PR 23-APR-2001; 2001US-0285749P.
 PR 03-MAY-2001; 2001US-0288327P.
 PR 03-MAY-2001; 2001US-0288504P.
 PR 29-MAY-2001; 2001US-0294047P.
 PR 30-MAY-2001; 2001US-0294473P.
 PR 08-JUN-2001; 2001US-0296964P.
 PR 18-JUN-2001; 2001US-0298959P.
 PR 19-JUN-2001; 2001US-0299324P.
 PR 13-AUG-2001; 2001US-0312026P.
 PR 16-AUG-2001; 2001US-0312889P.
 PR 21-AUG-2001; 2001US-0312908P.
 PR 28-AUG-2001; 2001US-0313390P.
 PR 31-AUG-2001; 2001US-0315470P.
 PR 07-SEP-2001; 2001US-0316447P.
 PR 07-SEP-2001; 2001US-0318115P.
 PR 12-SEP-2001; 2001US-0318118P.
 PR 19-SEP-2001; 2001US-0318740P.
 PR 18-OCT-2001; 2001US-0323379P.
 PR 18-OCT-2001; 2001US-0330245P.
 PR 18-OCT-2001; 2001US-0330308P.
 PR 14-NOV-2001; 2001US-0332701P.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Tchernev VT, Spytek KA, Zerhusen BD, Patturajan M, Shimkets RA;
 PI Li L, Gangolli EA, Padigaru M, Anderson DW, Rastelli L, Miller CE;
 PI Gerlach VL, Taupier RJ, Gusev VV, Colman SD, Wolenc AR, Pena CE;
 PI Furtak K, Grosse WM, Alsobrook JP, Lepley DM, Rieger DK, Burgess CE;
 XX WPI; 2002-706998/76.
 DR

XX New NOVX polypeptides and nucleic acids, useful for preventing or
PT treating NOVX-associated disorders, e.g. cancer, cardiomyopathy,
PT atherosclerosis, or diabetes, and in chromosome mapping, tissue typing or
PT pharmacogenomics.

XX Disclosure: SEQ ID NO 622; 1498pp; English.

XX This invention relates to a novel nucleic acids, and encoded polypeptides
XX thereof, which have properties related to the stimulation of biochemical
XX or physiological responses in a cell, tissue, organ or organism.
XX Specifically, it refers to the use of biologically active fragments for
XX diagnostic and prognostic assays and furthermore in the treatment of
XX diverse pathological conditions. The present invention describes novel
XX human and murine NOVX proteins, as well as methods to modulate their
XX expression using antisense oligos, ribozymes and peptide nucleic acids.
XX The NOVX polypeptides, polynucleotides and antibodies are useful in
XX treating or preventing NOVX-associated disorders, e.g. cardiomyopathy,
XX atherosclerosis, cancer and diabetes. Furthermore, they may be used in
XX treating or preventing diseases such as inflammation, autoimmune
XX disorders, allergies, blood disorders, acquired immunodeficiency syndrome
XX (AIDS), obesity, asthma, immunoglobulin (Ig) A nephropathy, cirrhosis,
XX arthritis, Alzheimer's disease, infections, stroke, muscular dystrophy
XX and epilepsy. Accordingly, these molecules have many activities including
XX cytoprotective, antioxidant, anti-inflammatory, immunosuppressive, antiallergic,
XX haemostatic, anti-HIV, antidiabetic, antiarteriosclerotic, anorectic,
XX antiasthmatic, nephrotropic, antiarthritic, hepatotropic,
XX neuroprotective, neurotropic, antibacterial, virucide, antiparasitic,
XX relaxant and anticonvulsant. In addition, they are useful in screening
XX assays to identify small molecules that modulate or inhibit, for example,
XX neurogenesis, wound healing and angiogenesis. The nucleic acids are also
XX used as in chromosome mapping, tissue typing, preventive medicine and
XX pharmacogenomics. This polypeptide is a homologue of a human NOVX protein
XX of the invention.

XX SQ Sequence 882 AA;

Alignment Scores:

Pred. No.: 4,266-248 Length: 882
Score: 2422.00 Matches: 465
Percent Similarity: 82.3% Conservatism: 0
Best Local Similarity: 82.3% Mismatches: 0
Query Match: 81.9% Indels: 100
DB: 5 Gaps: 1

US-10-825-632-6 (1-1669) x ADI117086 (1-882)

QY 2 ACAGGTACAGCAAAATCCTAAGTCACTTTTAAGATGTCAGAAATATGATGCTGAA 61
DB 318 ThrGlyThrAlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGlu 337
QY 62 GDAAGATCATAGATGTCATAGATAAGCACTAATTCACCTTTGAGATTCATTGAA 121
DB 338 GlyArgIleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGlu 357
QY 122 GDAAGTGAATATATGCGCAGAGCTGGATGAGTCTCTGAGGAGAAATATGCTGCTCATC 181
DB 358 GlyValGluTyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIle 377
QY 182 CTACTAGATCGCTCCAGACTCGCTCAGATAGTGTGATCTCACTGATTCATTTATC 241
DB 378 LeuLeuAspArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIle 397
QY 242 CAGTGAAGAGATGTTATGAAAGGAGAGACTCATTTGAGTCAGTCCCTGATTCCTGTG 301
DB 398 ProValGluAspValMetGluArgGlnArgLeuIleGluSerValProAspSerVal 417
QY 302 ACGCCAATAATATCTATGAAGAAACAACAGACATCTCGATAAATATCCATGACATCTTT 361
DB 418 ThrProLeuIleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePhe 437
QY 362 CATGTTTTTCCCAAGTCAAGAGGAAATTTGAGTTATTTTGGCTCTGATGCAAA 421

DB 438 HisValPheProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLys 457
QY 422 ACAGGTTTTCCGTCATTTATACAAAATTAATCATCTATTTTAAAGGAAAGCAATAATAACGA 481
DB 458 ThrGlyPheArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArg 477
QY 482 TCAGTGGTGGCTGCTGCTCCAGTCATTCAGTGTCTCAAGTGTCTCAAGAGGAGATAGCA 541
DB 478 SerSerGlyGlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAla 497
QY 542 ATTACAGTGGTGAATGGAAAGTTCTTGGCCGGCATCGATCATTAATATCCAAAGTTGATGA 601
DB 498 IleThrSerGlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGlu 517
QY 602 GTCAGAGGCTGGTATATTTTGAAGGCACCAAGACTCCCTTTAGAGCATCACTGTGAC 661
DB 518 ValArgArgLeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeuTyr 537
QY 662 GTAGTCAGTACGTAATCCTGGAGAGTGACAGGCTGACTGACCGTGGCTACTCACAT 721
DB 538 ValValSerTyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHis 557
QY 722 TCTTGTGTCATCAGTCAGCTGTGACTTCTTTATAAGTAAAGTATAGTAAACCAAGAAAT 781
DB 558 SerCysCysIleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsn 577
QY 782 CCACACTGTGTCTCCTTTACAAGCTAATAAGTCTCTGAAGATGACCACTGTCGCAAAACA 841
DB 578 ProHisCysValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThr 597
QY 842 AAGGAATTTTGGCCACCATTTTGGATTCAGCAGCTCTCTCTCTGACTACTACTCTCCA 901
DB 598 LysGluPheTrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProPro 617
QY 902 GAAATTTCTCTTTGAAAGTACTACTGATTTACATTTGATGGGATGCTCTCAAGCCT 961
DB 618 GluIlePheSerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysPro 637
QY 962 CATGATCTACAGCTGGAAAGAAATATCTACTGTGCTGTTTCATATATGTTGCTCTCAG 1021
DB 638 HisAspLeuGlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGln 657
QY 1021 ----- 1021
DB 658 ValGlnLeuValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAla 677
QY 1021 ----- 1021
DB 678 SerLeuGlyTyrValValValIleAspAsnArgGlySerCysHisArgGlyLeuLys 697
QY 1021 ----- 1021
DB 698 PheGluGlyAlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGly 717
QY 1021 ----- 1021
DB 718 LeuGlnTyrLeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGly 737
QY 1021 ----- 1021
DB 738 TrpSerTyrGlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArg 757
QY 1022 GTTGCTATGCTGGGCCCTCAGTCACTCTGTGGATCTCTATGATACAGGATACACGAA 1081
DB 758 ValAlaIleAlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGlu 777
QY 1082 CGTTATATGGTCCACCTCAGCAGATGAACAGGCTATTACTTAGGATCTGTGGCCATG 1141
DB 778 ArgTyrMetGlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMet 797
QY 1142 CAAGCAGAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGGTTCTCGAT 1201
DB 798 GlnAlaGluLysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAsp 817

QY 1202 GAGATGTCATTTGTCACATACCACTATATTACTAGTGTGAGGCTGGAAG 1261
Db 818 GUASnValHiePheAlaHisThrSerileLeuLeuSerPheLeuValArgAlaGlyLys 837
QY 1262 CCATATGATTACAGATCTATCTCTCAGGAGAGACACAGCATAGAGTTCTCTGAATCGGA 1321
Db 838 ProTyzAspLeuGlnIleTyzProGlnGluArgHisSerileArgValProGluSerGly 857
QY 1322 GAACATTATGAAGTGCATCTTTTGCACTACCTTCAAGAAACCTTGGATCACGTATTGCT 1381
Db 858 GluHisTyzGluLeuHisLeuLeuHisTyzLeuGlnGluLeuGlySerArgGlieAla 877
QY 1382 GCTCTAAAGTGATA 1396
Db 878 AlaLeuLysValIle 882

RESULT 14
ID ABU07720
AC ABU07720 standard; protein; 882 AA.
XX
DT 19-MAY-2003 (first entry)
XX
DE Human serine protease HIPHUM46.
XX
KW Human; enzyme; HIPHUM46; serine protease; gene therapy; osteoarthritis;
KW serine protease activity modulation; dipeptidyl peptidase activity;
KW musculoskeletal disease; Hepatitis B virus infection; myotonic dystrophy;
KW Alzheimer's disease; parasupranuclear palsy; Huntington's disease;
KW amyotrophic lateral sclerosis; malabsorption syndrome; lung disease;
KW irritable bowel syndrome; type I diabetes; faecal incontinence; anaemia;
KW haemorrhoid; proctitis; rectal polyp; small bowel tumour; dyslexia;
KW colorectal tumour; ceroid lipofuscinosis; allergic encephalomyelitis;
KW multiple sclerosis.
XX
OS Homo sapiens.

XX Key Location/Qualifiers
FH 259..260
FT Region /note= "Paired glutamates of the beta propeller domain"
FT Active-site 739
FT /label= Catalytic_serine_residue
FT Active-site 817
FT /label= Catalytic_aspartate_residue
FT Active-site 849
FT /label= Catalytic_histidine_residue
XX
PN GB2374869-A.
XX
XX 30-OCT-2002.
XX
XX 22-JAN-2002; 2002GB-00001404.
XX
XX 23-JAN-2001; 2001GB-00001760.
XX
XX (GLAX) GLAXO GROUP LTD.
XX
XX Edbrooke MR, Lewis AP;
XX WPI; 2003-150703/15.
XX N-PSDB; ABX12255.
XX
XX Identifying modulators of serine protease activity useful for treating
XX musculoskeletal diseases, by contacting cell expressing a novel serine
XX protease polypeptide with a compound and monitoring serine protease
XX activity.
XX
XX Claim 10; Page 26-29; 38pp; English.
XX
XX The invention relates to a method of identifying a substance that
XX modulates serine protease activity, comprising contacting a cell such as
XX

a neuronal cell, lung cell, intestinal cell or a cell infected with a virus, expressing a serine protease polypeptide (HIPHUM 46), or its variant having dipeptidyl peptidase activity, or a serine protease isolated from the cell with a test substance, and monitoring for serine protease activity. The method is useful for identifying a substance that modulates serine protease activity. A modulator of the serine protease is useful in the manufacture of a medicament for treatment or prophylaxis of a musculoskeletal disease e.g. osteoarthritis, Hepatitis B virus infection, Alzheimer's disease, parasupranuclear palsy, myotonic dystrophy, Huntington's disease or amyotrophic lateral sclerosis. Additional disease that may be treated using modulators of the serine protease include malabsorption syndromes, irritable bowel syndrome, lung disease, type I diabetes, faecal incontinence, haemorrhoids, proctitis, rectal polyps, small bowel tumours, colorectal tumours, anaemia, dyslexia, ceroid lipofuscinosis, allergic encephalomyelitis, and multiple sclerosis. The present sequence represents the amino acid sequence of the human serine protease HIPHUM46

XX
SQ Sequence 882 AA;
XX

Alignment Scores:
Pred. No.: 4,26e-248 Length: 882
Score: 2422.00 Matches: 465
Percent Similarity: 82.3% Conservative: 0
Best Local Similarity: 82.3% Mismatches: 0
Query Match: 81.9% Indels: 100
DB: Gaps: 1

US-10-825-632-6 (1-1669) x ABU07720 (1-882)

QY 2 ACAGGTACAGCAAACTCTAAAGTCACTTTTAAGATGTCAGAAATAATGATTGCTGAA 61
Db 318 ThrGlyThrAlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGlu 337
QY 62 GGAAGGATCATAGTGCATAGATAGCACTAATTCACCTTTTGAGATCTATTGAA 121
Db 338 GlyArgGlieIleAspValIleAspLysGluIleGlnProPheGluIleLeuPheGlu 357
QY 122 GGAGTTGATATATTGCCAGAGCTGGATGAGTCTCTGAGGAGAAATATGCTTGGTCAATC 181
Db 358 GlyValGluTyzIleAlaArgAlaGlyTrpThrProGluGlyLysTyzAlaIlePheSerIle 377
QY 182 CTACTAGATCGTCCAGACTCGCCTACAGATAGTGTGTGATCTCAGTGAATTTATTC 241
Db 378 LeuLeuAspArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIle 397
QY 242 CCAGTAGAAGATGATGTTATGGAAGGACAGAGACTCATTTGAGTCAGTGCCTGATTCTGTG 301
Db 398 ProValGluAspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerVal 417
QY 302 ACGCCACTAATTATCTATGAGAAACACAGACATCTGGATAAATATCCATGACATCTTT 361
Db 418 ThrProLeuIleIleTyzGluGluThrThrAspIleTrpIleAsnIleHisAspIlePhe 437
QY 362 CATGTTTTTCCCAAGTCACAGAGAAATGAGTTTATTTTTCCTCTGATGCAAA 421
Db 438 HisValPheProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLys 457
QY 422 ACAGGTTTTCCCTCATTTATACAAATFACATCTATTATAAGGAGAAATATTAACGA 481
Db 458 ThrGlyPheArgHisLeuTyzLysIleThrSerIleLeuLysGluSerLysTyzLysArg 477
QY 482 TCCAGTGTGGCTGCTGCTCCAGAGTATTTCAAGTGTCTTCAAGAGAGAGATGACA 541
Db 478 SerSerGlyGlyLeuProAlaProSerAspPheLysCysProIleLysGluIleAla 497
QY 542 ATTACCAAGTGTGAATGGAGTCTTGGCCGATGGATCTAATATTCCAAGTGAATGA 601
Db 498 IleThrSerGlyGluIlePheGluValLeuGlyArgHisGlySerAsnIleGlnValAspGlu 517
QY 602 CTCAGAGCTGCTGATATTTTGAAGGCCAACAGATCCCTTTTAGAGCATCACCTGTAC 661
Db 518 ValArgArgLeuValTyzPheGluGlyThrLysAspSerProLeuGluHisLeuTyz 537

QY 662 GTAGTCAGTTCAGTAAATCTCGAGAGGTGACAAAGGCTGACTGACCGTGGCTACTCACAT 721
DB 538 ValValSerTyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHis 557
QY 722 TCTTCTGCATCAGTCAGTCAGTCAGTCCTCTTTATAGTAAAGTATAGTAAACAGAGAAAT 781
DB 558 SerCysCysIleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsn 577
QY 782 CCACACTGTGTCTCCCTTTACAAGCTATCAAGTCTCAAGTCAAGCACTTGCAGAAACA 841
DB 578 ProHisCysValSerLeuTyrLysLeuSerSerProGluAspProThrCysLysThr 597
QY 842 AAGGAATTTTGGCCACCATTTTGGATTTCAGCAGGCTCTCTCTCTGACTATATCTCTCCA 901
DB 598 LysGluPheThrAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProPro 617
QY 902 GAAATTTCTCTTTTGAAGTACTCTGATTTACATTTATGATGGGATGCTCTACAGCCT 961
DB 618 GluIlePheSerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysPro 637
QY 962 CATGATCTACAGCTCGGAAGAAATATCCTACTGTGCTTCATATATGCTGCTCTCAG 1021
DB 638 HisAspLeuGlnProGlyLysTyrProThrValLeuPheIleTyrGlyGlyProGln 657
QY 1021 ----- 1021
DB 658 ValGlnLeuValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAla 677
QY 1021 ----- 1021
DB 678 SerLeuGlyTyrValValValValIleAspAsnArgGlySerCysHisArgGlyLeuLys 697
QY 1021 ----- 1021
DB 698 PheGluGlyAlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGly 717
QY 1021 ----- 1021
DB 718 LeuGlnTyrLeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGly 737
QY 1021 ----- 1021
DB 738 TrpSerTyrGlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArg 757
QY 1022 GTTGTATGTCTGGGGCCCCAGTCACCTCTGTGGATCTCTATGATACAGGATACACGGA 1081
DB 758 ValAlaIleAlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGlu 777
QY 1082 CATTATATGGTCCACCTGACCAAGTCAACAGGCTATTACTTAGGATCTGTGGCCATG 1141
DB 778 ArgTyrMetGlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMet 797
QY 1142 CAAGCAGAAAAGTCCCTCTGAACCAATCGTTTACTGTCTTACATGTTTCTCGAT 1201
DB 798 GlnAlaGlyLysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAsp 817
QY 1202 GAGAAATGTCATTTGTCATACACAGTATATCTAGTATTTTGTAGTGGCTGGAAG 1261
DB 818 GluAsnValHisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLys 837
QY 1262 CATTATGATTTACAGATCTATCTCTCAGGAGACACAGCATAGAGTTCTCTGAATCGGA 1321
DB 838 ProTyrAspLeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGly 857
QY 1322 GAACATTATGAACTGATCTTTTGCATCTACCTTCAAGAAAACCTTGGATCACCTATGCT 1381
DB 858 GluHisTyrGluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAla 877
QY 1382 GCTCTAAAAGTGATA 1396
DB 878 AlaLeuLysValIle 882

RESULT 15
ADR41398
ID ADR41398 standard; protein; 587 AA.
XX AC ADR41398;
XX 07-OCT-2004 (first entry)
XX DE Human CD-like molecule HDQBS11, SEQ ID NO:197.
XX Human; CD-like molecule; cluster of differentiation; diagnosis;
KW prevention; immune disorder; immunodeficiency; autoimmune disorder;
KW blood-related disorder; haematological disorder; haemostatic disorder;
KW thrombolytic disorder; hyperproliferative disorder; cancer; tumour;
KW apoptotic disorder; cardiovascular disorder; respiratory disorder;
KW angiogenic disorder; neovascularisation; neurological disorder;
KW endocrine disorder; reproductive system disorder; infectious disease;
KW gastrointestinal disorder; drug screening; tissue regeneration;
KW chemotaxis; gene therapy; antibody therapy; drug targeting;
KW chromosome mapping; forensic analysis; immunophenotyping; cytostatic;
KW haemostatic; tranquiliser; vulnerary; anti-inflammatory; nephrotropic;
KW cardiant; antiallergic; anti-HIV; antirheumatic; antiarthritic;
KW antipsoriatic; immunosuppressive; vasotropic; neuroprotective;
KW antithyroid; thyromimetic; gynaecological; virucide; hepatotropic;
KW antibacterial; dermatological; chromosome 15q22.
XX OS Homo sapiens.
XX PN WO200226930-A2.
XX PD 04-APR-2002.
XX PF 25-SEP-2001; 2001WO-US029838.
XX PR 26-SEP-2000; 2000US-0235484P.
XX FA (HUMA-) HUMAN GENOME SCI INC.
XX PI Rosen CA, Birse CE;
XX DR WPI; 2002-405050/43.
XX DR N-PSDB; ADR41222.
XX PT Novel polynucleotides and polypeptides useful for treating, preventing or
PT ameliorating cardiovascular, renal, neurovascular, and autoimmune
XX disorders.
XX FS Claim 11; SEQ ID NO 197; 1243pp; English.
XX CC The invention relates to 167 novel human CD (cluster of differentiation)-
CC like molecules (ADR41388-ADR41563) and to cDNAs encoding them (seqid:11)-
XX SQ Sequence 587 AA;
Alignment Scores:
Pred. No.: 5,07e-247 Length: 587
Score: 2411.00 Matches: 463
Percent Similarity: 82.2% Conservative: 0
Best Local Similarity: 82.2% Mismatches: 0
Query Match: 81.5% Indels: 100
DB: 5 Gaps: 1
US-10-825-632-6 (1-1669) x ADR41398 (1-587)
QY 8 ACAGCAATCTCTAACTCACTTTTAAAGTGTGAGAAATAATGATTGATGCTGAAGGAAGG 67
DB 25 ThrAlaAsnProLysValThrPheLysMetSerGluIleAspAlaGluGlyArg 44
QY 68 ATCATAGATCTCATAGATAAGGAACCTTAATTCACCTTTTGAGATTCTATTGAAGGAGTT 127
DB 45 IleIleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyVal 64
QY 128 GAATATATATCCAGAGCTGGATGCTCTCTGAGGAAAAATATGCTTGTCTCATCTACTA 187

```
Db 65 GluTyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeu 84
QY 188 GATCGCTCCCGACTCGCTCAGAGTAGTGTGATCTCACCTGAATATTATCCAGTA 247
Db 85 AspArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProVal 104
QY 248 GAAGATGATGTTATGAGAGGAGAGACTCATTGAGTCAGTCCTGATCTCTGACGCCA 307
Db 105 GluAspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrPro 124
QY 308 CTAATTATCTATGAGAAACACACACATCTGGATAAATATCCATGACATCTTTCATGT 367
Db 125 LeuIleIleTyrGluGluThrThrAspIleIlePheIleAsnIleHisAspIlePheHisVal 144
QY 368 TTTCCCAAGTCACGAGAGAAATTGAGTTTATTTTGCCTCTGAATCAAAACAGGT 427
Db 145 PheProGlnSerHisGluGluIleGluPheIlePheAlaSerGluCysLysThrGly 164
QY 428 TTCCGTCATTATACAAATATACATCTATTTTAAAGGAAAGCAAAATATAAACGATCCAGT 487
Db 165 PheArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSer 184
QY 488 GGTGGCTGCCTGCTCCAAGTGAATTTCAAGTGTCTCTATCAAGAGGAGATAGCAATTACC 547
Db 185 GlyGlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThr 204
QY 548 AGTGTGTAATGGAAAGTCTTGGCCGGCATGGATCTAATATCAAGTTGATGAAGTCAGA 607
Db 205 SetGlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArg 224
QY 608 AGGCTGTGATATTTTGAAGGACCAAGACTCCCTTTAGAGCATCACCTGTACGTAGTC 667
Db 225 ArgLeuValTyrPheGluGlyThrLysAspSerProLeuGluHisIleLeuTyrValVal 244
QY 668 AGTTAGTAAATCCTGGAGAGGTGACAGGCTGACTGACCGTGGCTACTCACAATCTTGC 727
Db 245 SerTyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCys 264
QY 728 TGCATCAGTCAGCACTGTGCTCTTTATAGTAAGTATAGTAACAGAGATCCACAC 787
Db 265 CysIleSerGlnHisCysAspPheIleSerLysTyrSerAsnGlnLysAsnProHis 284
QY 788 TGTGTGTCCTTTACAGCTATCAAGTCTCGAGATGACCCAACTTGCAGAAACAAAGGAA 847
Db 285 CysValSerLeuTyrLysLeuSerSerProGluAspSerProThrCysLysThrLysGlu 304
QY 848 TTTTGGCCACCATTTTGGATTACAGAGTCTCTTCCTGACTATATCTCTCCAGAAATT 907
Db 305 PheTrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIle 324
QY 908 TTCTCTTTTGAAGTACTGATTTACATTTGATGGGATGCTCTACAGCCTCATGAT 967
Db 325 PheSerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAsp 344
QY 968 CTACAGCTGGAAAGAAATATCTTACTGTCTGCTATATATGTTGGTCTCTCAG----- 1021
Db 345 LeuGlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGln 364
QY 1021 ----- 1021
Db 365 LeuValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeu 384
QY 1021 ----- 1021
Db 385 GlyTyrValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGlu 404
QY 1021 ----- 1021
Db 405 GlyAlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGln 424
QY 1021 ----- 1021
```

```
Db 425 TyrLeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSer 444
QY 1022 -----GTTGCT 1027
Db 445 TyrGlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAla 464
QY 1028 ATTGCTGGGCCCCACTCTCTGTGGATCTTCTTATGATACAGGATACACGGAACGTTAT 1087
Db 465 IleAlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyr 484
QY 1088 ATGGGTACCCCTGACCAGAAATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCA 1147
Db 485 MetGlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAla 504
QY 1148 GAAAAGTTCCCTCTCAACCAATCGTTTACTGTCTTACATGGTTTCTCGATGAGAAAT 1207
Db 505 GluLysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsn 524
QY 1208 GTCCATTTTGCACATACCAAGTATATTACTGAGTTTTTTAGTGAGGCTGGAAAGCCATAT 1267
Db 525 ValHisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyr 544
QY 1268 GATTTACAGATCTATCTCAGGAGAGACACAGCATTAAGAGTTCTCTGAATCGGAGAACAT 1327
Db 545 AspLeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHis 564
QY 1328 TATGAACTGCATCTTTTGCACCTACCTTCAAGAAACCTTGGATCAAGTATGCTGCTCTA 1387
Db 565 TyrGluLeuHisLeuLeuHisTyrLeuGlnGluLeuGlySerArgIleAlaAlaLeu 584
QY 1388 ARAAGTCATA 1396
Db 585 LysValIle 587
```

Search completed: May 2, 2006, 01:02:50
Job time : 290.372 secs

OM nucleic - protein search, using frame_plus_n2p model

Run on: May 2, 2006, 01:03:49 ; Search time 11.3267 Seconds
(without alignments)
2835.522 Million cell updates/sec

Title: US-10-825-632-6
Perfect score: 2957
Sequence: 1 aacagctacgcaaatccta.....aaaaaaaaaaaaaaaaaaaaa 1669

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 566832

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ n2p.model -DEV=xlp
-O=/abs/ABSSWEB.spool/US10825632/runat_01052006_105948_3262/app_query.fasta_1
-DB=PIR -QFMT=fastan -SUFFIX=n2p.rpr -MINMATCH=0.1 -LOOFC=0 -LOOPEXT=0
-UNITS=bits -STAR=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=pcpt -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs02p
-USER=US10825632 @CGN 1 1 92 @runat_01052006_105948_3262 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : PIR 80:.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match Length DB ID	Description
1	462	15.6 931 2 T32919	hypothetical prote
2	429.5	14.5 738 2 A87516	dipeptidyl peptida
3	355	12.0 741 2 JC5142	X-Pro dipeptidyl-p
4	300.5	10.2 711 2 S66261	X-Pro dipeptidyl-p
5	277	9.4 818 1 A30107	dipeptidyl aminope
6	267.5	9.0 931 2 A49737	dipeptidyl aminope
7	267	9.0 793 2 T41703	dipeptidyl aminope
8	248.5	8.4 792 1 A39914	dipeptidyl-peptida
9	238	8.0 766 1 CDHU26	dipeptidyl-peptida
10	230	7.8 760 1 S23752	dipeptidyl-peptida
11	226	7.6 779 2 T35173	hypothetical prote
12	226	7.6 799 2 T25174	hypothetical prote
13	221	7.5 803 2 A41793	dipeptidyl aminope
14	216	7.3 803 2 I68600	dipeptidyl aminope

15	216	7.3	865	2	I54331	dipeptidyl aminope
16	176	6.0	829	2	T19514	hypothetical prote
17	173.5	5.9	795	2	F82858	dipeptidyl-peptida
18	160.5	5.4	759	2	I38593	fibroblast activat
19	145.5	4.9	743	2	T37700	probable dipeptidyl
20	139	4.7	642	2	C71137	hypothetical prote
21	136.5	4.6	709	2	B82580	alanyl dipeptidyl
22	134	4.5	683	2	B87495	prolyl oligopeptid
23	131	4.4	631	2	H75007	hypothetical acylamino
24	113.5	3.8	466	2	B89877	hypothetical prote
25	111.5	3.8	1018	1	CGHU2A	collagen alpha 2(V
26	108.5	3.7	1415	2	S52267	DNA polymerase III
27	105	3.6	692	2	S61200	probable membrane
28	104.5	3.5	265	2	B84063	hypothetical prote
29	104.5	3.5	591	2	H72474	probable acylamino
30	104	3.5	417	2	JH0660	amine dehydrogenas
31	104	3.5	1438	2	C89900	DNA polymerase III
32	103.5	3.5	1835	2	S46082	urea carboxylase (
33	103.5	3.5	4452	1	YGBSG2	gramicidin S synth
34	102.5	3.5	592	2	C70108	peptidase homolog
35	102	3.4	732	1	A35655	peptidyl-dipeptida
36	102	3.4	1312	1	A34171	peptidyl-dipeptida
37	101.5	3.4	4450	2	JX0340	gramicidin S synth
38	101	3.4	1313	1	JC2038	serine C-palmitoyl
39	100	3.4	558	2	A43667	Probable acid-CoA
40	99.5	3.4	543	2	S46098	N conserved hypoch
41	99.5	3.4	1948	2	B69511	silencing protein
42	98.5	3.3	916	2	T51288	60 kDa inner membr
43	97	3.3	795	2	G72092	trab protein - Esc
44	97	3.3	866	2	I79267	hypothetical prote
45	96.5	3.3	440	2	T24073	

ALIGNMENTS

RESULT 1
T32919
hypothetical protein K02F2.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C;Accession: T32919
R;Maggi, L.; Goela, D.
submitted to the EMBL Data Library, January 1998
A;Description: The sequence of C. elegans cosmid K02F2.
A;Reference number: Z21246
A;Accession: T32919
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-931 <MAG>
A;Cross-references: UNIPROT:O4987; UNIPARC:UPI0000076BD8; EMBL:AF043699; PIDN:AAB97564
A;Experimental source: strain Bristol N2; clone K02F2
C;Genetics:
A;Gene: CESP:K02F2.1
A;Map position: 1
A;Introns: 58/3; 82/2; 131/2; 178/2; 275/3; 322/2; 404/2; 441/1; 464/2; 486/3; 528/2; 5

Alignment Scores:					
Pred. No.:	4.63e-31	Length:	931		
Score:	462.00	Matches:	154		
Percent Similarity:	38.7%	Conservative:	79		
Best Local Similarity:	25.6%	Mismatches:	177		
Query Match:	15.6%	Indels:	192		
DB:	2	Gaps:	23		
US-10-825-632-6 (1-1669) x T32919 (1-931) -					
Qy	5	GGTACAGCAAACTCTAAAGTCACCTTTTAAGATGTCAGAAATAATGATTGATGCTGAAGGA	64		
Db	355	GlyThrLysAsnAlaTyrSerThrLeuArgMet	-----ValIleLeuGluAsnGly	371	
Qy	65	AGGATCATGATGTCATAGATGAAGCAATAATCACTTTGATGATCTTATTGTAAGGA	124		
Db	372	LysAlaTyrAspValProLeuLysAsp	-----GluValIleTyrLysHis	386	

```
QY 125 -----GTTGAATATATTGCCAGAGCTGGATGACTCCTCGAGGGAATAATGCTTGG 175
Db 387 CysProPheTyrGluTyrIleThrArgAlaGlyPhePheSerAspGlyThrThrValTrp 406
QY 176 TCCATCTACTAGATCGCTCCAGACATCGCTACAGATAGTGTGATC----- 223
Db 407 ValGlnValMetSerArgAspGlnAlaGlnCysSerLeuLeuLeuIleProTyrThrAsp 426
QY 224 -----TCACCT 229
Db 427 PheLeuLeuProGluGluLeuGlyGlySerIleLysGluAspAenLeuGlnSerThr 446
QY 230 GAATATTATCCAGTAGAGATGATTATGGAAGCAGAGACTCATTCAGTCAGTG 289
Db 447 AspLeuAenMetGlyValTrpAspAsp---LysSerHisGluGluThrMetGluLysPro 465
QY 290 CTGATTCCTGTGACG-----CCACTAATTATCTTGAAGAACAACACACATCTGGATA 343
Db 466 ProArgGlyLysLeuArgGlyThrValGlnIleHisIleLysAlaArgAsnAspTyrTrpIle 485
QY 344 AATATCCATGATCTTTTCATGTTTTTCCCAAGATCAGCAAGAGGAA-----ATTGAG 397
Db 486 AenThrHisAenAlaIleTyrProLeuLysIleThrAspGluGluHisProMetTyrGlu 505
QY 398 TTTATTTTGGCTCTCGAATGCMAAACAGGTTTCGGTCATTATACAAAATTACATCTATT 457
Db 506 PheIleTyrCysLeuGluLysProAenGly---SerCysLeuAlaLeuIleSerAlaGlu 524
QY 458 TTAAGAGAAACAAATATAAACGATCCAGTGGTGGCTGCTGCTCCAGTAGTTTCAAG 517
Db 525 LeuAepGlnAenGlyTyr----- 530
QY 518 TGTCTCATCAAGAGGAGATAGCAATTTACCAGTGGTGATGGAGTCTTGTGGCCGCAT 577
Db 531 CysArgHisThrGluGluLysLeuLeuMetAlaGluAenPheSerIle-----AsnLys 548
QY 578 GGATCTAATATCCAAAGTTGATGAAGTCAGAGCGCTGTATATTTTGAAGCACCAAGAC 637
Db 549 SerMetGlyIleValValAspGluValArgGluLeuValTyrTyrValAlaAsnGluSer 568
QY 638 TCCCTTTAGACATCACCTGTAGTACGTAGTACGTAGTAAATCCTGGAGAGGTGACAAGG 697
Db 569 HisProThrGluTrpAenIle---CysValSerHisTyrArgThrGlyGlnHisAlaGln 587
QY 698 CTGACTACCGTGGTACTACATCTTCTGCTGCATCAGTCACACTGTGACTTCTTTATA 757
Db 588 LeuThrGluSerGlyIle-----Cys-----PheLys 596
QY 758 AGTAAGTATAGTAACCAAG-----AATCCACACTGTGTGTCTCCCTTTACAAG 805
Db 597 SerGluArgAlaAenGlyLysLeuAlaLeuAspLeuAspHisGlyPheAlaCysTyrMet 616
QY 806 CTATCA-----AGTCTGAAGATGACCCCAACTTGCAAAACAAAGGAATTT---TGG--- 853
Db 617 ThrSerValGlySerProAlaGlu-----CysArgPheTyrSerPheArgTrpLys 633
QY 854 -----GCCACCATTTTGGATTTCAGCAGGTCT 880
Db 634 GluAsnGluValLeuProSerThrValTyrAlaAlaAenIleThrValSerGlyHisPro 653
QY 881 ---CTTCCTGAC-----TATACTCTCCAGAAATTTTCTCTTTTGAAGT---ACTACT 928
Db 654 GlyGlnProAspLeuHisPheAspSerProGluMetIleGluPheGlnSerLysLysThr 673
QY 929 GGATTTACATTGATGGGATGCTTACAAAGCTCATGATCTACAGCCTCGGAAGAATAT 988
Db 674 GlyLeuMetHisTyrAlaMetIleLeuArgProSerAenPheAspProTyrLysLysTyr 693
QY 989 CCTACTGTGCTGTATATATGCTGCTCT----- 1018
Db 694 ProValPheHisTyrValTyrGlyGlyProGlyIleGlnIleValHisAenAspPheSer 713
```

```
QY 1018 ----- 1018
Db 714 TrpIleGlnTyrIleArgPheCysArgLeuGlyTyrValValValPheIleAspAenArg 733
QY 1018 ----- 1018
Db 734 GlySerAlaHisArgGlyIleGluPheGluArgHisIleHisLysLysMetGlyThrVal 753
QY 1018 ----- 1018
Db 754 GluValGluAspGlnValGluGlyLeuGlnMetLeuAlaGluArgThrGlyGlyPheMet 773
QY 1018 ----- 1018
Db 774 AspMetSerArgValValValHisGlyTrpSerTyrGlyTyrMetAlaLeuGlnMet 793
QY 1019 -----CAGTTTGCTATTGCTGGGCCCCCAGCTCACTCTGCTGG 1054
Db 794 IleAlaLysHisProAenIleTyrArgAlaAlaIleAlaGlyIleAlaValSerAspTrp 813
QY 1055 ATCTTCTATGATACAGGATACACGACGTTATATGGTCCACCTGACCAAGATGAACAG 1114
Db 814 ArgLeuTyrAspThrAlaTyrThrGluArgTyrMetGlyTyrPro---LeuGluGluHis 832
QY 1115 GCGTATTACTAGGATCTGTGGCCATCCAAAGCAGAAAAGTTCCCTCTGAAACCAATCGT 1174
Db 833 ValIlyGlyAlaSerSerIleThrGlyLeuValGluLysLeuProAspGluProAsnArg 852
QY 1175 TTACTGCTCTTACATGGTTTCTGGATGAGAATGTCCATTTTGCACATACCATATATTA 1234
Db 853 LeuMetLeuValHisGlyLeuMetAspGluAsnValHisPheAlaHisLeuThrHisLeu 872
QY 1235 CTGAGTTTTTTAGTAGGGCTGGAAAGCCATATGATTTTACAGATCTATCTCAGGAGAGA 1294
Db 873 ValAspGluCysIleLysLysGlyLysTyrPheIleGluLeuValIlePheProAsnGluArg 892
QY 1295 CACAGCATAGAGTCTCGAATCGGAGAAACATTTATGAACATGTCATCTTTTGACACTACCTT 1354
Db 893 HisGlyValArgAenAsnAlaSerIleTyrLeuAspAlaArgMetTyrPheAla 912
QY 1355 CAAGAA 1360
Db 913 GlnGln 914

RESULT 2
A:dipterydyl peptidase IV [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C>Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C:Accession: A87516
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Shapito, L.; Ventek, J.C.; Fraser, C.M
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249, MUID:21173698; PMID:11259647
A:Accession: A87516
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-738 <STO>
A:Cross-references: UNIPROT:Q9A6E0; UNIPARC:UPI000000C7616; GB:AE005673; NID:G13423647; F
C:Genetic: CC2154
A:Gene: C2154

Alignment Scores:
Pred. No.: 2,59e-28 Length: 738
Score: 429.50 Matches: 123
Percent Similarity: 36.5% Conservative: 57
Best Local Similarity: 24.9% Mismatches: 146
Query Match: 14.5% Indels: 167
DB: Gaps: 12

US-10-825-632-6 (1-1669) x A87516 (1-738)
```



```
QY 131 TATATCCAGAGCTGGATGACTCTCTGAGGAAATATATGCTGCTCATCTACTAGAT 190
Db 284 TyrValAlaArgValAlaTrpSerAlaAspGlyLysThrValTyrValGlnArgLeuSer 303
QY 191 CGCTCCAGAGCTCGCTCAGAGATGAGTGTGATCTCACCTGAATATTATTATCCAGTAGAA 250
Db 304 ArgAspGlnLysThrLeuAspLeuLeuAlaPhe----- 314
QY 251 GATGATGTTATGGAAGGCGAGAGACTCATTGAGTCAGTCGCTGATTCGTGACGCCACTA 310
Db 315 -----ATTATCTATGAAGAAACAACAGACATCTGATAAATATATCATGACTCTTTTCAT 364
QY 311 -----ATTTATCTATGAAGAAACAACAGACATCTGATAAATATATCATGACTCTTTTCAT 364
Db 321 GlyLysThrIleLeuThrAspThrAspProHisPheIleGluValSerAsnAspPheArg 340
QY 365 GTTTTTCCTCCAAAGTCACGAAGAGAAATTTGATTTTATTTTTCCTCTGAAATGCAAAACA 424
Db 341 ProlLeuThrAspGly-----ThrPheLeuTrpGlySerGlu---LysAsp 354
QY 425 GGTTCCTCGTCAATTTATACAAATATACATCTATTTTAAGGAAGCAATATATAACGATCC 484
Db 355 GlyAsnGlnHisLeuTyrArgTyrAla-----Ala 364
QY 485 AGTGGTGGCTGCTGCTCCAGTGATTTCAAGTGTCTATCAAGAGGAGGATAGCAATT 544
Db 365 AspGlyLysLeuIleAla-----GlnIle 372
QY 545 ACCAGTGTGTAATCGGAAGTTCTTTGGCCGCGATGATCTAATATCAAGTTGATCAAGTC 604
Db 373 ThrLysGlyAspTrpProValIleGlyLeuGluGly-----ValAspGluAla 388
QY 605 AGAAGGCTGGTATATTTTGAAGGACCACCAAGACTCCCTTTAGAGCATCAGCTGTACGTA 664
Db 389 ArgLysValAlaIlePheSerAlaSerIleAspThrProIleGluArgLeuTyrGlu 408
QY 665 GTCAGTTACGTAATCTCTGAGAGGTCACAAAGCTGACTGACCGTGGTACTCACATCT 724
Db 409 ValSerTyrAlaLysProGlyLysProLysAlaLeuThrSerAlaGlyGlyTrpTrpAla 428
QY 725 TGCTGCATCAGTCAGCACTGTGACTTCTTTTAAAGTAAGTATAGTAACCAAGAAATCCA 784
Db 429 AlaLysValAlaAspAsnGlyGlyAlaPheAlaGlyThrTyrSerAspProLysThrPro 448
QY 785 CACTGTGTGCTCTTACAGCTATCAAGTCCT-----GAAGTACACCCA 829
Db 449 SerGlnThrAlaLeuTyrSerAlaAspGlyLysArgValArgTrpIleGluGluAsnLys 468
QY 830 ACTTGCACAAACAAGGAATTTTGGCCACCATTGGAATTCAGAGGCTCTCTCTCTGCAC 889
Db 469 LeuAlaGluGlyHisProTyrTrp-----ProTyrAlaAlaAsnLeuProGln 484
QY 890 TATACTCTCCAGAAATTTCTCTTTTGAAGTACTGATTTACATTTGATGGGATG 949
Db 485 -----ProGluPheGlySerLeuLysAlaAspGlyGluThrLeuHisTyrGlu 501
QY 950 CTCTACAGCCTCATGATCTACAGCCTGGAAAGAAATATCCTACTGCTGTGTTTATATAT 1009
Db 502 IleLeuLysProIleGlyPheAspProAlaLysLysTyrProAlaIleValSerValTyr 521
QY 1010 GGTGTGCTCT----- 1018
Db 522 GlyGlyProHisAlaGlnArgValMetLysAsnTrpHisSerProSerGluArgThrTyr 541
QY 1018 ----- 1018
Db 542 LeuGluAlaGlyTyrValIlePheLysLeuAspAsnArgGlySerGlyAsnArgSerAla 561
QY 1018 ----- 1018
Db 562 LysPheMetArgAlaLeuAspArgLysLeuGlyThrValGluValGluAspGlnLeuLeu 581
```

```
QY 1018 ----- 1018
Db 582 GlyAlaLysPheLeuAlaSerGlnProTyrValAspAlaAspLysLeuGlyValMetGly 601
QY 1019 -----CAG 1021
Db 602 TrpSerTyrGlyGlyPheMetAlaLeuMetLeuLeuThrAlaGluAsnThrProPheLys 621
QY 1022 GTTCTCTATTGCTGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGAA 1081
Db 622 AlaGlyAlaAlaGlyAlaProProThrGluTrpSerLeuTyrAspThrAlaTyrThrGlu 641
QY 1082 CGTTATATGGGTCACTCCCTGACCAAGATGAACAGGGCTATTACTAGGATCTGTGCCCATG 1141
Db 642 ArgTyrMetGlyLysProAspGluAsnLysAlaGlyTyrAlaTyrSerAspIleAsnAsn 661
QY 1142 CAAGCAGAAAGTTCCCTCTGAACCAAAATCGTTTACTCTCTTACATGTTTCTCGAT 1201
Db 662 ArgIleAspLysLeu-----AlaProGlySerLeuLeuLeuLeuHisGlyMetAlaAsp 679
QY 1202 GAGATGTCCATTTTGCACATACCACTATATTAAGTGTGAGTGTGAGTGTGAGGCTCGAAAG 1261
Db 680 AsnAsnValIlePheGluAsnSerThrArgLeuMetAlaLeuGlnArgLysAlaIle 699
QY 1262 CCATATGATTTACAGATCTATCTCTCAGGAGAGACACAGC 1300
Db 700 LeuPheGluMetAlaMetTyrProGlyGluArgHisSer 712
RESULT 3
JC5142
X-Pro dipeptidyl-peptidase (EC 3.4.14.11) - Xanthomonas maltophilia
C:Species: Xanthomonas maltophilia
C>Date: 31-Jan-1997 #sequence_revision 27-Feb-1997 #text_change 09-Jul-2004
C:Accession: JC5142
R:Kabashima, T.; Ito, K.; Yoshimoto, T.
J. Biochem. 120, 1111-1117, 1996
J. Title: Dipeptidyl peptidase IV from Xanthomonas maltophilia: Sequencing and expression
A:Reference number: JC5142; MUID: 97164011; PMID: 9010758
A:Accession: JC5142
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-741 <KAB>
A:Cross-references: UNIPROT:P95782; UNIPARC:UPI0000085F8; DDBJ:D83263; NID:gl753196; PI
C:Comment: This enzyme catalyzes the hydrolytic removal of amino-terminal dipeptidyl re:
C:Superfamily: dipeptidyl-peptidase IV
C:Keywords: dipeptidyl-peptidase, membrane bound
F:4-18/Domain: transmembrane #status predicted <TMM>
F:610/Active site: Ser #status predicted
F:685/Active site: Asp #status predicted
F:717/Active site: His #status predicted
Alignment Scores:
Pred. No.: 5, 93e-22 Length: 741
Score: 355.00 Matches: 123
Percent Similarity: 36.1% Conservative: 77
Best Local Similarity: 22.2% Mismatches: 164
Query Match: 12.0% Indels: 190
DB: 2 Gaps: 17
US-10-825-632-6 (1-1669) x JC5142 (1-741)
QY 5 GGTACAGCAAAATCTTAAAGTCACCTTTTAAGATGTCAGAAATATGATGATGCTGAAGA 64
Db 257 GlyAspAlaAsnValGlnValLysLeuGlyValIleSerProAlaGluGlnAlaGlnThr 276
QY 65 AGATCATAGATGTCATAGATAAGCACTAATCAACCTTTTGAGATTCCTATTGTAAGGA 124
Db 277 GlnTrpIleAspLeuGlyLysGluGlnAspIle----- 287
QY 125 GTTGAATATATTGCCAGAGCTGGAGCTCCTCAGGAGAAATATGCTTGGTCCATCCTTA 184
Db 288 -----TyrLeuAlaArgValAsnTrpArgAspProGlnHisLeuSerPheGln---Arg 304
```



```
Db      355 AspTrpGluIleLeuLeuAsnTyrGlyTyrAsnProLysThrLysGlu----- 370
QY      614 GTATATTTTGAAGCACCAGAACTCCCTTTTAGACATCACTGTACGTAGTCAGT--- 670
Db      371 ValTyrIleGlnThrThrGluLysGlySerIleAsnLys-----ValValSerLys 387
QY      671 ---TACGTAATCTCGAGAGGTGACAAAGCGTCACAGCGTCCGCTACTCACATCTTGC 727
Db      388 LeuAsnIleAsnThrGlyLysThrGlnLeuLeuSerAsnAlaGluGlyAsnAsnSerAla 407
QY      728 TGCATCAGTCAGCAGCTGTGACTCTTTTATAAGTAAGTATAGTAACCCAGAGAATCCACAC 787
Db      408 AlaPheSerLysThrPheAsnTyrPheIleAsnThrSerSerThrAlaLysValProThr 427
QY      788 TGTGTGTCCTTTTACAAGCTATCAAGTCCTGAA-----CATGAC 826
Db      428 LysTyrIleLeuLysAspAlaAsnGlyLysAspValLysGluLeuGlnAsnAsnAspAsp 447
QY      827 CCACTTCGAAACAAAGGAATTTTGGGCCACCATTTTGGATTTCAGCAGGTCTCTTCT 886
Db      448 LeuLeuAsnLysLeuLysSer-----Asp 455
QY      887 GACTATATCTCTCCAGAAATTTCTCTTTTGAAGTACTACTGGATTACATTGATGGG 946
Db      456 AsnPheIleAlaLysGluPheIleThrIleProAsnAlaAlaGlyAspGlnMetAsnAla 475
QY      947 ATGCTCTACAAGCTCATGATCTACAGCTCGGAAAGAAATATCTCTGCTGCTTCATA 1006
Db      476 TrpMetIleLysProLysAsnPheAspProAlaLysLysTyrProValPheMetPheGln 495
QY      1007 TATGTGTGTCCT-----CAGTTGCT----- 1027
Db      496 TyrSerGlyProGlySerGlnGlnValAlaAsnSerTrpAspGlyGlyAsnGlyIleTrp 515
QY      1027 ----- 1027
Db      516 PheAspMetLeuAlaGlnLysGlyTyrLeuValValCysValAspGlyArgGlyThrGly 535
QY      1027 ----- 1027
Db      536 PheArgGlyThrLysTyrLysLysValThrTyrLysAsnLeuGlyLysTyrGluIleGlu 555
QY      1027 ----- 1027
Db      556 AspGlnIleThrAlaAlaLysTrpLeuGlyAsnGlnSerTyrValAspLysSerArgIle 575
QY      1027 ----- 1027
Db      576 GlyIlePheGlyTrpSerTyrGlyTyrMetAlaSerLeuAlaMetThrLysGlyAla 595
QY      1028 -----ATTGCTGGGGCCCCAGTCACCTGTGGATCTCTTATGATACA 1069
Db      596 AspValPheLysMetGlyIleAlaValAlaProValThrAsnTrpArgPheTyrAspSer 615
QY      1070 GGATACAGGAACGTTATATGGGTCACCTGACAGATGAACAGGGCTATTACTATGA 1129
Db      616 IleTyrThrGluArgPheLeuGlnThrProGlnGluAsnLysAspGlyTyrAspLeuAsn 635
QY      1130 TCTGTGGCCATCGAAGCAGAAAGTTCCCTCTGACCAANTGTTTACTGCTCTTACAT 1189
Db      636 SerProThrThrTyrAlaLysLeuLeu-----LysGlyLysPheLeuLeuIleHis 652
QY      1190 GGTTCCTCGATGAGATGTCATTTCGACATACAGATATATTACTGAGTTTATTAGTG 1249
Db      653 GlyThrAlaAspAspAsnValHisPheGlnAsnSerMetGluPheSerGluAlaLeuIle 672
QY      1250 AGCGCTGGAAAGCCATATGATTATACAGATCTATCTCTCAGGACAGACACAGCATAGAGTT 1309
Db      673 GlnAsnLysLysGlnPheAspPheMetAlaTyrProAspLysAsnHisSerIleIleGly 692
QY      1310 CCTGAATCGGAGAACATTATGAACATGCTCTTTTGGCACTACCTTCAAGAAACCTT 1366
Db      693 GlyAsnThrArgProGlnLeuTyrGluLysMetThrAsnTyrIleLeuGluAsnLeu 711
```

RESULT 5

```
A30107
dipeptidyl aminopeptidase B (EC 3.4.14.-) - yeast (Saccharomyces cerevisiae)
N;Alternate names: protein YHR028c
C;Species: Saccharomyces cerevisiae
C;Date: 07-Jun-1990 #sequence_revision 30-May-1997 #text_change 09-Jul-2004
C;Accession: S46780; A30107
R;Db, Z.
submitted to the EMBL Data Library, June 1994
A;Description: The sequence of S. cerevisiae cosmid 8082.
A;Reference number: S46773
A;Accession: S46780
A;Molecule type: DNA
A;Residues: 1-818 <DUZ>
A;Cross-references: UNIPROT:P18962; UNIPARC:UPI0000031ASF; EMBL:U10399; NID:G500689; PII
R;Roberts, C.J.; Pohlrig, G.; Rothman, J.H.; Stevens, T.H.
J. Cell Biol. 108, 1363-1373, 1989
A;Title: Structure, biosynthesis, and localization of dipeptidyl aminopeptidase B, an i
A;Reference number: A30107; MUID:89174971; PMID:2647766
A;Accession: A30107
A;Molecule type: DNA
A;Residues: 1-82, 'H', 84-124, 'N', 126-181, 'LRRLLET', 189-199, 'N', 201-365, 'DFKRGKERKF', 376-5
A;Cross-references: UNIPARC:UPI0000172A3C; EMBL:X15484
A;Note: the authors translated the codon ACC for residue 572 as Asn
C;Genetics:
A;Gene: SGD:DAP2; STE13; MIPS:YHR028c
A;Cross-references: SGD:S0001070; MIPS:YHR028c
A;Map position: 8R
C;Superfamily: dipeptidyl-peptidase IV
C;Keywords: dipeptidylpeptide hydrolase; glycoprotein; transmembrane protein; yeast vac
F;30-45/Domain: transmembrane #status predicted <TM>
F;63,79,110,139,392,421/Binding site: carbohydrate (Asn) (covalent) #status predicted

Alignment Scores:
Pred. No.:      2,77e-15      Length:      818
Score:          277.00      Matches:      121
Percent Similarity: 32.2%      Conservative: 64
Best Local Similarity: 21.0%      Mismatches: 170
Query Match:      9.4%      Indels:      220
DB:              1          Gaps:      22

US-10-825-632-6 (1-1669) x A30107 (1-818)
QY      143 GCTGATGACTCCTCGAGGGAAATATGCTTGGTCCATCTACTAGATCGCTCCAGACT 202
Db      251 AlaTrpTrpSerProThrGlyAspTyrLeuAlaPheLeuLysIleAspGluSerGluVal 270
QY      203 CGCTACAGATAGTCTTCATCTCACCTCGAATATTATTATCCAGTAGAAGATGATGTATG 262
Db      271 GlyGluPheIleIle-----ProTyrTyrValGlnAspGluLysAspIleTyr 286
QY      263 GAAAGCGACAGACTCATTTAGTCAGTCGCTGATTCCTGTGACGCCACTTAATTATCTATGAA 322
Db      287 ProGluMetArgSerIleLys---TyrProLysSerGlyThrPro-----Asn 301
QY      323 GAAACACACAGACATCTGGATA---AATATCCATGAC-----ATCTTTGATGTTTCCC 373
Db      302 ProHisAlaGluLeuTrpValTyrSerMetLysAspGlyThrSerPheHisProArgIle 321
QY      374 CAAAGTCACGAAGAGAAATTGAGTTATATTTT----- 406
Db      322 SerGlyAsnLysLysAspGlySerLeuLeuIleThrGluValThrTrpValGlyAsnGly 341
QY      407 -----GCCTCTGAATGCAAAACAGGTTTCCGTCATTTATACAAAATTACA 451
Db      342 AsnValLeuValIleThrThrAspArgSerSerAspIleLeuThrValPheLeuIleAsp 361
QY      452 TCTATTATTTAAAGGAAGCAAAATATAACCA-----TCCAGTGGTGGG----- 493
Db      362 ThrIleAlaLysThrSerAsnValValArgAsnGluSerSerAsnGlyGlyTrpTrpGlu 381
QY      494 -----CTGCCTGCTCCAAAGTATTCAAGTGCTCCTATCATCAAGAG 532
```


Alignment Scores:

Pred. No.: 1,86e-14 Length: 931
Score: 267.50 Matches: 114
Percent Similarity: 33.0% Conservative: 63
Best Local Similarity: 21.3% Mismatches: 164
Query Match: 9.0% Indels: 195
DB: 2 Gaps: 18

US-10-825-632-6 (1-1669) x A49737 (1-931)

```
QY 221 ATCTCACTGAAATTTATTCACAGTAGAAGATGATTTATGGAAGCAGAGACTCATT 280
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 445 ILeSerProAspThrPhe-----ArgPheGluIleThrAspArgAsnSerLysIle 461
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 281 GAATCAGTGCCT-----GATTCGTGACGCCACTAATATCTATGAAGAACA 328
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 462 LeuAspValLysValTyAspIleProSerGlnMetLeuThrValArgAsnThrAsn 481
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 329 ACAGACATC-----TGATA-----AATATCCATGACATCTTTCATGTTTCCCA 376
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 482 SerAsnLeuPheAsnGlyTrpIleGluLysThrLysAspIleLeuSerIleProProLys 501
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 377 AGTCACGAAGGAATTAG-----TTTATTTTGGCTCTGATGCAAAACAGGTTTC 430
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 502 ProGluLeuLysArgMetAspTyGlyTyIleAspIleHisAlaAspSerArgGlyPhe 521
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 431 CGTCATTTATACAAATATCATCTATTTTAAAGGAAGCAATATATACAGTCCAGTGGT 490
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 522 SerHisLeuPheTyTyProThrValPhe----- 531
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 491 GGGCTGCTGCTCCAGTGATTTCAAGTGTCTATCAAGAGAGAGATAGCAATACCAGT 550
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 532 -----AlaLysGluProIleGlnLeuThrLys 540
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 551 GGTGAATGGGAAGTTCTTGGCGGCAT-----GGATCTAAATATCAAGTTGATGAAGTC 604
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 541 GlyAsnTrpGluValThrGlyAsnGlyIleValGlyTyTrGluTyGluThrAsp----- 558
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 605 AGAAGGTGGTATATTTGAAGGCACCAAGACTCCCTTTAGACATCACTGTACGTA 664
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 559 -----ThrIlePheThrAlaAsnGluIleGlyValMetSerGlnHisLeuTySer 576
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 665 GTCAGTTACGTA-----AATCCTGGAGAG 688
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 577 IleSerLeuThrAspSerThrThrGlnAsnThrPheGlnSerLeuGlnAsnPro----- 594
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 689 GTGACAGGCTGACTGACCTGGCTACTCACAATCTTGTGCTCAGTCAGCACTGTGAC 748
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 595 -----SerAspLysTyAspPheTyAspPheGluLeuSerSerSerAlaArg 610
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 749 TTCCTTTAAGTAGTAGTAGTAACCAAGAAATCCA----- 784
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 611 TyrAlaIleSerLysLysLeuGlyProAspThrProIleLysValAlaGlyProLeuThr 630
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 785 -----CACTGTGTCCTCTTACAAAGCTATCAAGTCTGA 820
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 631 ArgValLeuAsnValAlaGluIleHisAspSerIleLeuGlnLeuThrLysAspGlu 650
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 821 GATGACCCCACTTCGAAACAAAGAAATTTTGGGCCACCATTTTGGATTACAGGTTCT 880
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 651 -----LysPheLysGlu-----Lys 655
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 881 CTTCTGACTACTCTCCAGAAATTTTCTCTTTTGAAGTACTACT-----GGA 931
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 656 IleLysAsnTyAspLeuPro-----IleThrSerTyLysThrMetValLeuAspAspGly 674
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 932 TTATCATTTGATGGATGCTCTCAAGCCCTCATGATCTACAGCTCGGAAGAAATATCCT 991
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 675 ValGluIleAsnTyIleGluIlePheProAlaAsnLeuAsnProLysLysTyPro 694
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 992 ACTGTGCTGTTTCATATATGGTGGTCT----- 1018
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 695 IleLeuValAsnIleTyGlyGlyProGlySerGlnThrPheThrThrLysSerSerLeu 714
```

```
QY 1018 ----- 1018
Db 715 AlaPheGluGlnAlaValSerGlyLeuAspValIleValLeuGlnIleGluProArg 734
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1018 ----- 1018
Db 735 GlyThrGlyLysGlyTyTrpSerPheArgSerTipAlaArgGluLysLeuGlyTyTrp 754
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1018 ----- 1018
Db 755 GluProArgAspIleThrGluValThrLysLysPheIleGlnArgAsnSerGlnHisIle 774
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1018 ----- 1018
Db 775 AspGluSerLysIleAlaIleTrpGlyTyTrpSerTyGlyGlyPheThrSerLeuLysThr 794
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1019 -----CAGGTGCTATTGCTGGGGCCCCCAGCTCACTCTG 1051
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 795 ValGluLeuAspAsnGlyAspThrPheLysTyAlaMetAlaValAlaProValThrAsn 814
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1052 TGGATCTTCTATGATACAGGATACACGGAACGTTATATGGTCAACCCTGACCAAGTAA 1111
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 815 TrpThrLeuTyAspSerValTyThrGluArgTyTrpMetAsnGlnProSerGluAsnHis 834
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1112 CAGGGCTATTACTTAGATCTGTGGCATGCAAGCAGAAAAGTTCCCTCTGACCAATAAT 1171
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 835 GluGlyTyPheGluValSerThrIleGlnAsnPheLysSerPheGluSer---LeuLys 853
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1172 CGTTTACTGCTCTTACATGGTTTCTGGATGAGAAATGTCATTTTGCATACCATCAGTATA 1231
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 854 ArgLeuPheIleValHisGlyThrPheAspAspAsnValHisIleGlnAsnThrPheArg 873
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1232 TTACTGAGTTTTTGTAGTGAGGGCTGA---AAGCCATATGATTATACAGATCTATCTCAG 1288
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 874 LeuValAspGlnLeuAsnLeuLeuGlyLeuThrAsnTyAspMetHisIlePheProAsp 893
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1289 GAGAGACACACATAGAGTTCTCGAATCGGAGAACATTTATGAATGTCATCTTTTGCAC 1348
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 894 SerAspHisSerIleArgTyHisAsnAlaGlnArgIleValPheGlnLysLeuTyTrp 913
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1349 TACCTTCAAGAAAACCTTGGATCAGTCATGCTGCTCTATAAAGTGATA 1396
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 914 TrpLeuArgAspAlaPheAlaGluArgPheAspAsnThrGluValLeu 929
```

RESULT 7

T41703
dipeptidyl aminopeptidase - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: T41703
R;Murphy, L.; Harris, D.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
submitted to the EMBL Data Library, August 1998
A;Reference number: Z22011
A;Accession: T41703
A;Status: preliminary; translated from GB/EMBL/DDBU
A;Molecule type: DNA
A;Residues: 1-793 <MUR>
A;Cross-references: UNIPROT:O14073; UNIPARC:UPI000013AB95; EMBL:AL031180; PIDN:CAA20138
A;Experimental source: strain 972h-; cosmid c2811 -chimeric
C;Genetics:
A;Gene: SPAC2E11.08
A;Map position: 1
C;Superfamily: dipeptidyl-peptidase IV

Alignment Scores:
Pred. No.: 1,96e-14 Length: 793
Score: 267.00 Matches: 100
Percent Similarity: 35.1% Conservative: 37
Best Local Similarity: 25.6% Mismatches: 127
Query Match: 9.0% Indels: 126
DB: 2 Gaps: 14

C;Comment: This protein is localized to the bile canaliculus, which is the apical domain
 C;Superfamily: dipeptidyl-peptidase IV
 F;1-792/Product: dipeptidyl-peptidase, membrane-bound form #status experimental <SIG>
 F;1-28/Domain: signal sequence #link MATS #status experimental <MATM>
 F;1-6/Domain: intracellular #status predicted <INT>
 F;7-28/Domain: transmembrane #status predicted <TMN>
 F;29-792/Domain: extracellular #status predicted <EXT>
 F;29-34/Domain: propeptide #link MATS #status experimental <PRO>
 F;35-792/Product: dipeptidyl-peptidase, soluble form #status experimental <MATS>
 F;83,90,148,217,227,319,521,686/Binding site: carboxylate (Asn) (covalent) #status predicted
 F;631/Active site: Ser #status experimental
 F;709,741/Active site: Asp, His #status predicted

Alignment Scores: 7.45e-13 Length: 792
 Pred. No.: 248.50 Matches: 122
 Score: 32.4% Conservative: 66
 Percent Similarity: 32.4% Mismatches: 188
 Best Local Similarity: 21.0% Indels: 205
 Query Match: 8.4% Gaps: 23
 DB: 1

US-10-825-632-6 (1-1669) x A39914 (1-792)

Qy 5 GGTAACGCAAACTCTAAAGTCACCTTTTAAGATGTCCAGAAATAATGATTGATGCTGAAGGA 64
 Db 258 GYAlaValAsnProThrValLysPhePhele-----ValAsnThrAspSerLeuSer 275
 Qy 65 AGGATCATAGATCATAGATAGGAACATAATCAACCTTTTGAGATTCTATTGAAGGA 124
 Db 276 SerThrThrThrProMetGlnleThrAlaProAlaSerValThr-----ThrGly 294
 Qy 125 GTTGAATATATCCAGAGCTGGATGGACTCTCTGAGGGAATAATGCTTGTGTCATCCTA 184
 Db 295 AspHisTyrLeuCysAspValAlaTrpValSerGluAspArgIleSer----- 310
 Qy 185 CTAGATCGCTCCAGAGCTCGCCTACAGATAGTGTGTGATCTCACCTGAATTATTATCCCA 244
 Db 310 ----- 310
 Qy 245 GTAGAGATGATGTTATGAAAGGCGAGAGACTCATGATGAGTCCCTGATTCGTGACG 304
 Db 311 -----LeuGlnTrpLeuArgArgIleGlnAsnTyr-----SerValMet 323
 Qy 305 CCATAATTATCATGAGAAACACAGACATCTGATAATAATATCATGACATCTTTTAT 364
 Db 324 AlaIleCysAspTyrAspLysThrAsnLeuValTrpAsnCysProThrThrGlnGluHis 343
 Qy 365 GTT-----TTTCCCAAAAGTCAC 382
 Db 344 IleGluThrSerAlaThrGlyTrpValGlyArgPheArgProAlaGluProHisPheThr 363
 Qy 383 GAAGAGAAATGAGTTTATTTTGGCTCTGAATGCAAAACAGGTTTCCGTCAATTATAC 442
 Db 364 SerAspGlySerPheTyrLysIleValSerAspLysAspGlyTyrLysHis----- 381
 Qy 443 AAATTACATCTATTTAAGGAAACCAATATAACGATCCAGTGGGGCTGCCTGCT 502
 Db 382 ---IleCysGlnPheGlnLysAspArgLys----- 390
 Qy 503 CCAAGTGATTTCAAGTGTCTCTATCAAGAGGAGATAGCAATATACCAGTGGTGAATGGAA 562
 Db 391 -----ProGluGlnAspCysThrPheIleThrLysGlyAlaTrpGlu 404
 Qy 563 GTTCTTGGCCGGATGATCATATATCAAGTTGATGAAGTCAGAAAGCTCGGTATATTT 622
 Db 405 ValIle-----SerIleGluAlaLeuThrSerAspTyrLeuTyrTyrIle 419
 Qy 623 GAAGGCACC---AAAGACTCCCTTTAGAGCATCACCTGATAGTACGTTACGTAAAT 679
 Db 420 SerAsnGluTyrLysGluMetProGlyArgAsnLeuTyrLysIle----- 435
 Qy 680 CTGGAGAGGTGACAAAGGCTGACTGACCGGTGGCTACTCACATTCCTTGCTGCATCAGT--- 736

Db 436 -----GlnLeuThrAsp-----HisThrAsnLysLysCysLeuSerCys 448
 Qy 737 -----CAGCACTGTGACATCTCTTT-----ATAAGTAAGTATAGTAACACAG 775
 Db 449 AspLeuAsnProGluArgCysGlnTyrTyrSerValSerLeuSerLysGluAlaLysTyr 468
 Qy 776 AAGAAATCCACACTGT-----GTGTCCCTTTTCAAGCTATCAAGTCTCGAAGAT 823
 Db 469 TyrGlnLeuGlyCysArgGlyProGlyLeuProLeuTyrThrLeuHisArgSerThrAsp 488
 Qy 824 GACCAACTTGCACAAACAAGAAATTTGGGCCACCATTTGGATTCACAGAGTCCCTCTT 883
 Db 489 GlnLysGluLeuArgValLeuGluAsp---AsnSerAlaLeuAspLysMet-----Leu 505
 Qy 884 CCGTACTATATCTCTCCAGAA-----ATTTCCTCTTTTGAAGTACTACTGGAATTT 934
 Db 506 GlnAspValGlnMetProSerLysLysLeuAspPheIleValLeuAsnGluThrArgPhe 525
 Qy 935 ACATTGTATGGGATGCTCTACAGCCTCATGATCTACAGCCTGGAAAGAAATATCTCTACT 994
 Db 526 ---TriPtyrGlnMetIleLeuProProHis---PheAspLysSerLysLysTyrProLeu 543
 Qy 995 GTGCTGTTCATATATGCTGCTCT----- 1018
 Db 544 LeuIleAspValTyrAlaGlyProCysSerGlnLysAlaAspAlaAlaPheArgPheAsn 563
 Qy 1018 ----- 1018
 Db 564 TrpAlaThrTyrLeuAlaSerThrGluAsnIleIleValAlaSerPheAspGlyArgGly 583
 Qy 1018 ----- 1018
 Db 584 SerGlyTyrGlnGlyAspLysIleMetHisAlaIleAsnLysArgLeuGlyThrLeuGlu 603
 Qy 1018 ----- 1018
 Db 604 ValGluAspGlnIleGluAlaAlaArgGlnPheLeuLysMetGlyPheValAspSerLys 623
 Qy 1019 CAGTCTCTCTATGCTGGG----- 1036
 Db 624 GlnValAlaIleTrpGlyTrpSerTyrGlyTyrValThrSerMetValLeuGlySer 643
 Qy 1037 -----GCCCCAGTCACTCTGTGGATCTTCTAT 1063
 Db 644 GlySerGlyValPheLysCysGlyIleAlaValAlaProValSerArgTrpGluTyrTyr 663
 Qy 1064 GATACAGATACACGNAACGTTATATGGGT-----CACCTGACCAAGTTCGACAGGCG 1117
 Db 664 AspSerValTyrThrGluArgTyrMetGlyLeuProThrProGluAspAsnLeuAspHis 683
 Qy 1118 TATTACTTAGTCTGTGGCCATGCAAGCAAGAAAGTTCCTCTCTGAACCAAAATCGTTTA 1177
 Db 684 TyrArgAsnSerThrValMetSerArgAlaGluAsnPhe-----LysGlnValGluTyr 701
 Qy 1178 CTGCTCTTACATGTTTCTCGATGAGATGTCATTTGTCACATACACGATATATTATCTG 1237
 Db 702 LeuLeuIleHisGlyThrAlaAspAspAsnValHisPheGlnGlnSerAlaGlnIleSer 721
 Qy 1238 AGTTTTTTAGTGGGCTGGAAGCCATATGATTTTACAGATCTATCTCTCAGGAGAGACAC 1297
 Db 722 LysAlaLeuValAspAlaGlyValAspPheGlnAlaMetTyrTyrThrAspGluAspHis 741
 Qy 1298 AGCATTAAGAGTCTCTGAATCGGGAACATATTGAACATCTCATCTTTTGCACCTCAAC 1357
 Db 742 GlyIleAlaSerSerThrAlaHisGlnHisIleTyrSerHisMetSerHisPheLeuGln 761
 Qy 1358 GAA 1360
 Db 762 Gln 762
 Qy 1358 GAA 1360
 Db 762 Gln 762

RESULT 9
 CDH26

	...			:		
498	SefAlaLeuAspLysMet-----	LeuGlnAsnValGlnMetProSerLysLysLeuAsp	515			
QY	TTCCTCTTTGAAGACTACCTGGATTTTACATTGTATGGCATCTTACAAAGCCTCATGAT	967				
DB						
516	PheIleIleLeuAsnGluThrLysPhe--	TrpTyrGlnMetIleLeuProHis---	533			
QY	CPACAGCCTGGAAAATAATCTCTACTGTGCTCTTTCATATATGGTGTCCT-	1018				
DB						
534	PheAspLysSerLysLysTyrProLeuLeuLeuAspValTyrAlaGlyProCysSerGln	553				
QY	-----	-----	1018			
DB	LysAlaAspThrValPheArgLeuAsnTrpAlaThrTyrLeuAlaSerThrGluAsnIle	573				
QY	-----	-----	1018			
DB	IleValAlaSerPheAspGlyArgGlySerGlyTyrGlnGlyAspLysIleMetHisAla	593				
QY	-----	-----	1018			
DB	IleAsnArgArgLeuGlyThrPheGluValGluAspGlnIleGluAlaAlaArgGlnPhe	613				
QY	-----	-----	1018			
DB	SerLysMetGlyPheValaspAsnLysArgIleAlaIleTrpGlyTrpSerTyrGlyGly	633				
QY	-----	---CAGTTGCTATTGCTGGG	1036			
DB	TyrValThrSerMetValLeuglySerGlySerGlyValPheLysCysGlyIleAlaVal	653				
QY	GCCCCAGTCATCTGTGGATCTCTATGATACAGGATACACGGAACGTTATATGGGT---	1093				
DB	AlaProValSerArgTrpGluTyrTyrAspSerValTyrThrGluArgTyrMetGlyLeu	673				
QY	---CACCTGACCAGAATGAACAGGCCTATTACTTAGGATCTGTGGCCATCGCAAGCGAA	1150				
DB	ProThrProGluaspAsnLeuAspHisTyrArgAsnSerThrValMetSerArgAlaGlu	693				
QY	ANGTTCCTCTGAACCAATCGTTTACTCTCTTACATGGTTTCTCGATGACAATGTC	1210				
DB	AsnPhe-----	LysGlnValGluTyrLeuLeuIleHisGlyThrAlaAspAspAsnVal	711			
QY	CATTTTGCACATACCAGTATATTACTAGTTTTTTAGTCAGGCCGCGAAGCCATATGAT	1270				
DB	HisPheGlnGlnSerAlaGlnIleSerLysAlaLeuValaspValGlyValaspPheGln	731				
QY	TTTACAGATCTATCTCAGGAGACACAGCATTAAGAGTTCTCTGAATCGGGAGAACATTAT	1330				
DB	AlaMetTrpTyrThrAspGluaspHisGlyIleAlaSerSerThrAlaHisGlnHisIle	751				
QY	GAATGCGATCTTTGGCACTACCTTCAAGAA	1360				
DB	TyrThrHisMetSerHisPheIleLysGln	761				

RESULT 10

**KR3001
S23752**

S23732 dipeptidyl-peptidase IV (EC 3.4.14.5) alpha chain - mouse

N:Alternate names: CD26 alpha subunit; THAM alpha subunit
 arpeptidyl-peptidase IV (EC3.4.14.3) alpha chain - mouse

N; Alternate names: CD26 alpha subunit
C: Species: Mus musculus (house mouse)

C;species: MUB musculus (house mouse)
C;Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text change 10-Sep-1999

C;Date: 10-Sep-1999 #sequence_revis
C;Accession: S23752: A46465: A56030

C;Accession: S23/52; A46465; A56030
R:Marcuet D.: Bernard A M.: Vivier T.: Darmoul D.: Naciet P.: pierres. M.

R; Marguet, D.; Bernadotte, J. Biol. Chem. 267 22

J. Biol. Chem. 267, 2200-2208, 1992

A;Title: cDNA cloning for mouse thymocyte-activating mo-
A;Reference number: S23752;MIID:92130288;PWID:1370813

A;Reference number: S233752

A;Accession: S23752

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-760 <MAR>

A;Cross-references: UNIPARC:UPI0000172A2D; EMBL:X58384

R; Vivier, I.; Marguet, D.; Naq

A,Title: Evidence that thymocyte-activating molecule is mouse CD26 (dipeptidyl peptidase A)
A,Reference number: A46465; MUID:91302787; PMID:1712807
A,Accession: A46465
A,Status: preliminary
A,Molecule type: protein
A,Residues: 1-20 <VIV>
A,Cross-references: UNIPARC:UPI000172A2E
A,Experimental source: M14.T thymoma cells, Swiss nu/nu
A,Note: sequence extracted from NCBI backbone (NCBIP:42236)
B,Bernard, A.M.; Mattei, M.G.; Pierres, M.; Marguet, D.
Biochemistry 33, 15204-15214, 1994
A,Title: Structure of the mouse dipeptidyl peptidase IV (CD26) gene.
A,Reference number: A56030; MUID:95092780; PMID:7999781

Alignment Scores:

Argument Scores:	
Pred. No.:	2.8e-11
Score:	230.00
Percent Similarity:	33.5%
Best Local Similarity:	22.6%
Query Match:	7.8%
DB:	1
Length:	760
Matches:	98
Conservative:	47
Mismatches:	120
Indels:	168
Gaps:	19

US-10-825-632-6 (1-1669) x S23752 (1-760)

Qy	419	AAACACAGGTTTC	CGTCATTTTATACAAATTACATCTATTTTAAAGGAAGCAAAATATAA	478
			::: :::	
Db	372	LysAspGlyTyrLys	HisIleCysHisPhe	381
Qy	479	CGATCCAGTGTGGGCTCGCTCGTCCAAAGTGAATTTCAAG	--TGTCTATCAAGAGGAG	535
Db	382	-----ProlysAspLysLysAspCysThrPhe	-----	390
Qy	536	ATAGCAATTACCACTGGTGAATGGGAATTTCTTGGCCGGCATGGATCTAATATCCAAAGTT	595	
			::: :::	
Db	391	-----IleThrLysGlyAlaTrpGluValIle	-----SerIleGluAla	403
Qy	596	GATGAAGTCAGAAGGCTGGTATATTTTGAAGGCACC	---AAAGACTCCCTTTAGAGCAT	652
Db	404	LeuThrSerAspTyrLeuTyrTyrIleSerAsnGlnTyrLysGluMetProGlyGlyArg	423	
Qy	653	CACCTGTACGTAGTCAGTTACGTAAATCCTGGAGAGGTGACAAAGGCTGACTGACCGTGGC	712	
Db	424	AsnLeuTyrLysIle	-----GlnLeuThrAsp	432
Qy	713	TACTCACATCTTGTGTCATCAGT	-----CAGCACTGTGACTTCTTTATA	757
		::: :::		
Db	433	HisThrAsnValLysCysLeuSerCysAspLeuAsnProGluArgCysGlnTyrTyrAla	452	
Qy	758	AGTAAGTATATGTAACCAAGAAATCCAC	-----TGT	796
Db	453	ValSerPheSerLysGluAlaLysTyrTyrGlnLeuGlyCysTrpGlyProGlyLeuPro	472	
Qy	797	CTTTTCAAGCTATCAAGTCCCTGAAGATGACCCAACTTGGCAAAACAAAGAAATTTTGGGC	856	
Db	473	LeuTyrThrLeuHisIleArgSerThrAspHisLysGluLeuArgValLeuGlu	-----	489
Qy	857	ACCAATTTGGATTACGACAGGT	-----CCTCTCTGACTATATCTCTCCAGAA	904
		::: :::		
Db	490	-----AspAsnSerAlaLeuAspArgMetLeuGlnAspValGlnMetProSerLys	506	
Qy	905	-----ATTTTCTCTTTTGAAGTACTGCTGGATTTTACATTTGATGGGATGCTCTCAAG	958	

Db 507 LysLeuaspPheIleValLeuAsnGluThrArgPhe---TrpTyrGlnMetIleLeuPro 525
QY 959 CCTATGATCTACAGCTCGAAGAAATATCTCTACTGCTGCTCTGTTTCATATATGGTGGCTCT 1018
Db 526 ProHis---PheAspLysSerLysLysTyrProLeuLeuLeuAspValTyrAlaGlyPro 544
QY 1018 ----- 1018
Db 545 CysSerGlnLysAlaAspAlaSerPheArgLeuAsnTrpAlaThrTyrLeuAlaSerThr 564
QY 1018 ----- 1018
Db 565 GluAsnIleIleValAlaSerPheAspGlyArgGlySerGlyTyrGlnGlyAspLysIle 584
QY 1018 ----- 1018
Db 585 MetHisAlaIleAsnArgAlaGLeuGlyThrLeuGluValGluAspGlnIleGluAlaAla 604
QY 1019 ----- 1036
Db 605 ArgGlnPheValLysMetGlyPheValAspSerLysArgValAlaIleTrpGlyTrpSer 624
QY 1036 ----- 1036
Db 625 TyrGlyGlyTyrValThrSerMetValLeuGlySerGlySerGlyValPheLysCysGly 644
QY 1037 -----GCCAGTCACCTCTGCGATCTCTATGATACAGGATACAGGACGATAT 1087
Db 645 IleAlaValAlaProValSerArgTrpGluTyrTyrAspSerValTyrTrpGluArgTyr 664
QY 1088 ATGGGT-----CACCTGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCCATG 1141
Db 665 MetGlyLeuProIleProGluLeuAspAsnLeuAspHisTyrArgAsnSerThrValMetSer 684
QY 1142 CAAGCAGAAAGTTCCCTCTGAACCAAAATCGTTTACTGCTCTTACATGGTTCTCTGGAT 1201
Db 685 ArgAlaGluHisPhe-----LysGlnValGluTyrLeuLeuIleHisGlyThrAlaAsp 702
QY 1202 GAGATGTCCATTTTGACATACACAGTATATCTAGTGTGTTTATAGTGAGGCGCTGGAAG 1261
Db 703 AspAsnValHisPheGlnGlnSerAlaGlnIleSerLysValLeuValAspAlaGlyVal 722
QY 1262 CCATATGATTTACAGATCTATCTCTCAGGAGACACAGCATAAAGATTCTCTGAATCGGA 1321
Db 723 AspPheGlnAlaMetTrpTyrThrAspGluAspHisGlyIleAlaSerSerThrAlaHis 742
QY 1322 GAACATTATGAACCTGCACTTTTGCATCTACCTCTCAAGAA 1360
Db 743 GlnHisIleTyrSerHisMetSerHisPheLeuGlnGln 755
RESULT 11
T25173
hypothetical protein T23F1.7a - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T25173
R:Wilkinson, J.
submitted to the EMBL Data Library, October 1996
A:Reference number: Z19990
A:Accession: T25173
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-779 <Wtl>
A:Cross-references: UNIPROT:O18119; UNIPARC:UPI000002A220; EMBL:Z81129; PIDN:CAB03411.1
A:Experimental source: clone T23F1
C:Genetics:
A:Gene: CESP:T23F1.7a
A:Map position: 5
A:Introns: 13/3; 52/1; 111/3; 151/2; 177/3; 346/3; 504/1; 537/1; 628/2; 683/1; 723/1
C:Superfamily: dipeptidyl-peptidase IV
Alignment Scores:
Pred. No.: 6.18e-11 Length: 779

Score: 226.00 Matches: 99
Percent Similarity: 34.3% Conservative: 62
Best Local Similarity: 21.1% Mismatches: 142
Query Match: 7.6% Indels: 166
DB: 2 Gaps: 18
US-10-825-632-6 (1-1669) x T25173 (1-779)
QY 317 TATGAAGAAACACACAGCATCTGGATAAATATATCCATGACATCTTTTCATCTTTTCCCCAA 376
Db 350 TyrLysTyrAlaSerLysArgTrpValThr---HisAspPheHisSerIle---Thr 367
QY 377 AGTCACCAAGAGGAAATTTAGTTTATTTTTCCTCTGAAATGCAAAACAGGTTTTCGGTCAT 436
Db 368 SerPheGluAspThrLeuPhePheLeuLeuPro-----His 379
QY 437 TTATACAAAATATACATCTATTTTAAAGGAAAGCAAAATATAAACGATCCAGTCGTGGGTG 496
Db 380 AspLysArgAspAsnAlaPheGlnValAlaSerLeuArgLeuSerHisGlyGlnLeu 399
QY 497 CTGCTCCCAAGTGATTTTCAAGTGTCTTATCAAGAGGAGATAGCAATTACCAGTGGTGAA 556
Db 400 ArgThrProLys-----PheLeuAsnLeuGlyGlu 409
QY 557 TGGGAAGTTCTTGGCCGGCATGATCTAATATCCAAAGTTGATCAAGTCAGAGGCTGTA 616
Db 410 TyrAspValThrSerIleAsnGlyIleAsnLysGluThr-----ArgThrIle 425
QY 617 TATTTTGAAGGCCAACAAAGCTCCCTTTAGAGCATCAGCTGATGATGATGATGATGATGATG 676
Db 426 PhePheHisAlaAlaProLysProSerHisArgSerLeuPheSerTyrSer----- 443
QY 677 AATCTGTGAGAGGTGACACAGCGTACTGACCGGTGGTCTACTACATTTCTTGTGTCATGAT 736
Db 444 -----LeuAlaAspGluSerArgAsnSerAlaTyrCysIleSer 456
QY 737 -----CAGCACTCTGACTCTTTTATAAGTAAGTATAGTAAACCAAGAAATCCACAC 787
Db 457 CysSerIleLysAsnCysThrTrpAlaGlnAlaGlnMetAspAspGlnMetLysThrAla 476
QY 788 TGTGTGTCCTTTTACAAGCTATCAAGTCTCTGAAAGAT----- 823
Db 477 IleValSerCysLysGlyProAlaAlaProHisThrAlaIleValAsnLeuThrArgMet 496
QY 824 GACCCAACTTCAAAACAAAG-----GAATTTTGGGCCCACTT 862
Db 497 AspSerAspLysThrGluHisAlaAsnLeuLysAspLysThrTyrGlnAsnArg 516
QY 863 TTGATTACAGCAGGTCTCTCTGATATATCTCTCCAGAAATTTCTCTTTTCAAGAT 922
Db 517 ValGluGluAlaGly---LeuProValIleIleLysGluThrIle-----LysIle 532
QY 923 ACTACTGGATTTACATTTGATGGATGCTCTACAAGCTCATGATCATGATCATGATGATGATG 982
Db 533 SerAspAspPheAspAlaLeuIleLysLeuSerIleProLysAspIleTyrAsnArgAsp 552
QY 983 AATAT-----CCTACTGTGTTTCATATATGTTGTTCTCT----- 1018
Db 553 LysHisGlnAlaIleProLeuIleValHisValTyrGlyGlyProAsnAspGlnAsnThr 572
QY 1018 ----- 1018
Db 573 LysGluAlaThrGlnIleGlyIleGluGluValValAlaSerAlaSerGlnAlaIle 592
QY 1018 ----- 1018
Db 593 LeuArgIleAspGlyArgGlySerGlyArgGlyTrpLysTyrArgSerAlaIleTyr 612
QY 1018 ----- 1018
Db 613 GlyGlnLeuGlyThrValGluValGluAspGlnIleLysAlaIleLysValValLeuArg 632
QY 1019 -----CAGTTGCTATTGCTGGG----- 1036

```
Db 633 LeuTyrArgHisLeuLeuAspAlaArgValAlaValPheGlyTyrSerTyrGlyGly 652
QY 1036 -----
Db 653 PheMetThrLeuSerMetValAsnGluAlaProGluGlnPhePheLysCysAlaValSer 672
QY 1037 ---GCCCCAGTCACTCTGGGATCTTATGATACAGGATACACGGAACTTATATGGGT 1093
Db 673 ValAlaProValThrAsnPheAlaTyrThrAspAlaThrTyrThrGluAlaGlyMetGly 692
QY 1094 CACCTGACAGAAAGAACAGGCTATTACTTAGATCTGGCCATCGAAGCAGAAAG 1153
Db 693 -----AspAlaProLeuGluSerTyr-----SerAspValThrLysLysLeuAspAsn 708
QY 1154 TTCCCTCTGAACCAATCGTTTACTGCTTACATGCTTCTTCTGATGAGATGCCAT 1213
Db 709 PheLysSer-----ThrArgLeuLeuLeuMetHisGlyLeuLeuAspAsnValHis 726
QY 1214 TTTCACATACCACTATATCTAGTGTCTTTAGTGGGCTGGAAGCCATATGATTTA 1273
Db 727 PheGlnAsnSerAlaLeuLeuLeuAspGluLeuGlnAsnArgGlyValAspPheAspLeu 746
QY 1274 CAGATCTATCTCAGGAGACACAGCATTAAGAGTTCCTGAATCGGAGAACATTTAGAA 1333
Db 747 MetValTyrProAsnGlnAlaHisSerLeuSerSerArgThrSer-----HisValVal 764
QY 1334 CTGCATCTTTGACCTACCTTCAGAA 1360
Db 765 GlyLysMetThrHisPheLeuArgGln 773
RESULT 12
T25174
Hypothetical protein T25174 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T25174
R:Wilkinson, J
submitted to the EMBL Data Library, October 1996
A:Reference number: Z19990
A:Accession: T25174
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-799 <WtL>
A:Cross-references: UNIPROT:O18119; UNIPARC:UPI0000061190; EMBL:Z81129; PIDN: CAB03412.1
A:Experimental source: clone t23f1
C:Genetics:
A:Gene: CESP:T23f1.7b
A:Map position: 5
A:Introns: 13/3; 52/1; 111/3; 197/3; 366/3; 524/1; 557/1; 648/2; 703/1; 743/1
C:Superfamily: dipeptidyl-peptidase IV
Alignment Scores:
Pred. No.: 6.23e-11 Length: 799
Score: 226.00 Matches: 99
Percent Similarity: 34.3% Conservative: 62
Best Local Similarity: 21.1% Mismatches: 142
Query Match: 7.6% Indels: 166
DB: 2 Gaps: 18
US-10-825-632-6 (1-1669) x T25174 (1-799)
QY 317 TATGAAGAAACACAGACATCTCGATAAATATCCATGATCTTTCATGTTTTCGCCAA 376
Db 370 TyrLysTyrAlaSerLysArgTyrValThr-----HisAspAspPheHisSerIle---Thr 387
QY 377 AGTCACGAGAGGAATTTAGTTTATTTTTCCTCTGATGCAAAACAGGTTTCGTCAT 436
Db 388 SerPheGluAspThrLeuPhePheLeuPro-----His 399
QY 437 TTATACAAATTTACATCTATTTTAAAGGAAAGCAATATAACGATCCAGTGGTGGGCTG 496
Db 400 AspLysArgAspAsnAlaPheGlnValAlaSerLeuArgLeuSerHisGlyGlnLeu 419
```

```
QY 497 CCTGCTCCAAGTGATTTCAAGTGTCTATCAAAAGAGGAGATAGCAATTACCAGTGGTGA 556
Db 420 ArgThrProLys-----PheLeuAsnLeuGlyGlu 429
QY 557 TGGGAAGTCTTGGCGGCATGATCTAATATCAAGTTGATGAAGTCAGAGGCTGGTA 616
Db 430 TyrAspValThrSerIleAsnGlyIleAsnLysGluThr-----ArgThrIle 445
QY 617 TATTTTGAAGGACCAAGACTCCCTTTAGACATCACCTGTAGCTAGTACGTACGTACGTA 676
Db 446 PhePheHisAlaAlaAlaProLysProSerHisArgSerLeuPheSerTyrSer----- 463
QY 677 AATCCTGGAGGTGCAAGGCTGACTGACCTGGCTACTCATTCTTGTGCTGATCAGT 736
Db 464 -----LeuAlaAspGluSerArgAsnSerAlaTyrCysIleSer 476
QY 737 -----CAGCACTGTGACTTCTTATAGTATAGTAGTAAACCAAGAAATCCACAC 787
Db 477 CysSerIleLysAsnCysThrTrpAlaGlnAlaGlnMetAspAspGlnMetLysThrAla 496
QY 788 TGTGTGTCCTTTACAAGCTATCAAGTCTCCTGAAGAT----- 823
Db 497 IleValSerCysLysGlyProAlaAlaProHisThrAlaIleValAsnLeuThrArgMet 516
QY 824 GACCAACTTGCACAAACAAAG-----GAATTTTGGGCCACCCATT 862
Db 517 AspSerAspLysLysThrGluHisAlaAsnLeuLeuTyrAspLysThrTyrGlnAsnArg 536
QY 863 TTGATTACAGCAGTCTCTTCTGCTGATATCTCTCCAGAAATTTCTCTTTTGAAGT 922
Db 537 ValGluGluAlaGly---LeuProValIleIleLysGluThrIle-----LysIle 552
QY 923 ACTACTGGATTACATTTGATGGATGCTCTACAGCCTCATGATCTACAGCTGGAAG 982
Db 553 SerAspAspPheAspAlaLeuLysLeuSerIleProLysAspIleTyrAsnArgAsp 572
QY 983 AATATAT-----CCTACTGTGCTTCAATATATGTTGGTCTCT----- 1018
Db 573 LysHisGlnAlaIleProLeuIleValHisValTyrGlyGlyProAsnAspGlnAsnThr 592
QY 1018 ----- 1018
Db 593 LysGluAlaThrGlnIleGlyIleGluGluValAlaValAlaSerGlnAlaIle 612
QY 1018 ----- 1018
Db 613 LeuArgIleAspGlyArgGlySerGlyGlyArgGlyTyrLysTyrArgSerAlaIleTyr 632
QY 1018 ----- 1018
Db 633 GlyGlnLeuGlyThrValGluValGluAspGlnIleLysAlaIleLysValValLeuArg 652
QY 1019 -----CAGGTGTCTATTGCTGGG----- 1036
Db 653 LeuTyrArgHisLeuLeuAspAlaArgArgValAlaValPheGlyTyrSerTyrGlyGly 672
QY 1036 ----- 1036
Db 673 PheMetThrLeuSerMetValAsnGluAlaProGluGlnPhePheLysCysAlaValSer 692
QY 1037 ---GCCCCAGTCACTCTGATGATCTTCTATGATACAGGATACACGGAACTTATATGGGT 1093
Db 693 ValAlaProValThrAsnPheAlaTyrTyrAspAlaThrTyrThrGluArgTyrMetGly 712
QY 1094 CACCTGACAGAAAGAACAGGCTATTACTTAGGATCTGTGCCATCGAAGCAGAAAG 1153
Db 713 -----AspAlaProLeuGluSerTyr-----SerAspValThrLysLysLeuAspAsn 728
QY 1154 TTCCCTCTGAACCAATCGTTTACTGCTTACATGCTTCTTCTGATGAGATGCCAT 1213
Db 729 PheLysSer-----ThrArgLeuLeuLeuMetHisGlyLeuLeuAspAsnValHis 746
```

```
QY 1214 TTTCACATACCATATATCTAGTCTTTTGTAGTGGCTGGAAAGCCCATATGATTAA 1273
||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 747 PheGlnAenSerAlaIleLeuIleAspGluLeuGlnAenArgGlyValAspPheAspLeu 766
QY 1274 CAGATCTATCTCTCAGGAGACACACACATAAGAGTCTCGAATCGGAGAACATATGAA 1333
||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 767 MetValTyrProAenGlnAlaHisSerLeuSerSerArgThrSer-----HisValVal 784
QY 1334 CTGCATCTTTTGTGACCTACCTTCAAGAA 1360
||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 785 GlyLysMetThrHisPheLeuArgGln 793

RESULT 13
A:1793
dipeptidyl aminopeptidase-like protein (EC 3.4.14.-) short form, DPPX-S - bovine
C:Species: Bos primigenius taurus (cattle)
C>Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C:Accession: A41793
R:Wada, K.; Yokotani, N.; Hunter, C.; Doi, K.; Wentholt, R.J.; Shimasaki, S.
Proc. Natl. Acad. Sci. U.S.A. 89, 197-201, 1992
A:Title: Differential expression of two distinct forms of mRNA encoding members of a dip
A:Reference number: A41793; MUID:92108018; PMID:1729689
A:Accession: A41793
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-803 <WAD>
A:Cross-references: UNIPROT:P42659; UNIPARC:UPI00002A83C; GB:M76429; NID:94087119; PIDN:
A:Note: sequence extracted from NCBI backbone (NCBI:P75136)
A:Superfamily: dipeptidyl-peptidase IV
C:Keywords: dipeptidylpeptide hydrolase; glycoprotein
F:257,342/Binding site: carbohydrate (Aen) (covalent) #status predicted

Alignment Scores:
Pred. No.: 1,67e-10 Length: 803
Score: 221.00 Matches: 107
Percent Similarity: 32.8% Conservative: 65
Best Local Similarity: 20.4% Mismatches: 178
Query Match: 7.5% Indels: 174
DB: 2 Gaps: 19

US-10-825-632-6 (1-1669) x A41793 (1-803)
QY 131 TATATTCCAGAGCTGGATGGACTCTCTGAGGAAATATGCT-----TGGTCCATCCTTA 184
||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 320 TyrIleThrMetValLysTrpAlaThrSerThrLysValAlaValAenTrp----- 336
QY 185 CTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACTGAAATTTATATCCCA 244
||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 337 LeuSerArgAlaGlnAen-----ValSer----- 344
QY 245 GTAAAGATGATGTTATGAAAGCGAGACTATTGAGTCAGTCCTGATTCTGTGACG 304
||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 345 -----TATGAAGAAACAACAGACATCTGGATAAATATCCATGACATC 353
QY 305 CCACTAATTATC-----TATGAAGAAACAACAGACATCTGGATAAATATCCATGACATC 358
||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 354 GlyValCysThrLysLysHisGluAspGluSerGluAlaTrpLeu----- 368
QY 359 TTTTCATGTTTTTCCCAAGTCCACAGAGGAATTCAGTTTATTTTTCCTCTGATGC 418
||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 369 -----HisArgGlnAenGluProValPheSer----- 378
QY 419 AAAACAGGTTTCGTCATTATACAAATTCATCTATTAAAG-----GAAACCAA 472
||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 379 LysAspGly--ArgLysPhePheValArgAlaIleProGlnGlyGlnGlyLys 397
QY 473 TATAAAGCATCAGTGGTGGCTGCTCTCAAGTGATTTCAAGTGCTCTATCAAGAG 532
||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 398 PheTyrHisIleThrValSerSerSerGlnProAenSerSerAenAspAenIleGln--- 416
QY 533 GAGATAGCAATTCACAGTGGTGAATGGAGTCTTCTGGCGGATGATCTAATATCCAA 592
||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 417 -----SerIleThrSerGlyAspTrpAspValThr-----LysIleLeuSer 430
```

```
QY 593 GTTGATGAAGTCACAGGCTGGTATATTTTGAAGGCACCAAAAGACTCCCTTTAGAGCAT 652
||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 431 TyrAspGluLysArgSerGlnIleTyrPheLeuSerThrGluAspLeuProArgArgArg 450
QY 653 CACCTGACGTAGTCAGTCTAGTAAATCTCTGGAGAGGTGCACAAG----- 697
||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 451 GlnLeuTyrSerAlaSerThrVal-----GlySerPheAsnArgGlnCysLeuSerCys 468
||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 698 ---CTGACTGACCGTGGCTACTCACAATCTTCTGTCATCAGTCAGCACTGTGACTCTTTT 754
||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 469 AspLeuValAspAsnCysThrTyrPheSerAlaSerPheSerProGlyAlaAspPhePhe 488
QY 755 ATAAGTAAAGTATAGTAACAGAGAATCCACACTGTGTCTCTTCTTACAGCATATCAAGT 814
||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 489 LeuLeuLysCysGluGlyProGlyValPro---ThrValSerValHisAenThrThrAsp 507
QY 815 CCTCAAGTACGCCCACTTGCACAAACAAAGAA---TTTGGGCCCACTTTTGGATTCA 871
||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 508 LysLysLysMetPheAspLeuGluThrAsnGluHisValGlnLysAlaIleSerAspArg 527
QY 872 GCAGGTCTCTCTCTGACTATCTCTCCAGAAATTTCTCTTTTGAAGTACTACTGGA 931
||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 528 GlnMetProLysValGluTyrArgLysIleGlu-----ThrAspAsp 541
QY 932 TTTACATGTTGGGATGCTCTACAGCCTCATGATCTACAGCTGGGAAGAAATATCTCT 991
||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 542 TyrAsnLeuProIleGlnIleLeuLysProAlaThrPheThrAspThrAlaHisTyrPro 561
QY 992 ACTGTGCTGTTCATATATGCTGGTCTCT----- 1018
||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 562 LeuLeuLeuValValAspGlyThrProGlySerGlnSerValAlaGluLysPheAlaVal 581
QY 1018 ----- 1018
Db 582 ThrTrpGluThrValMetValSerSerHisGlyAlaValValLysCysAspGlyArg 601
QY 1018 ----- 1018
Db 602 GlySerGlyPheGlnGlyThrArgLeuLeuHisGluValArgArgLeuGlySerLeu 621
QY 1018 ----- 1018
Db 622 GluGluLysAspGlnMetGluAlaValArgValMetLeuLysGluProTyrIleAspLys 641
QY 1019 ---CAGGTTGCTATTGCTGGG----- 1036
||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 642 ThrArgValAlaValPheGlyLysAspTyrGlyGlyTyrLeuSerThrTyrLeuLeuPro 661
QY 1037 -----GCCCCAGTCACT 1048
Db 662 AlaLysGlyAspGlyGlnAlaProValPheSerCysGlySerAlaLeuSerProIleThr 681
QY 1049 CTGTGGATCTTCTATGATACAGATACACGAGAGCTTATATGGTCCACCTGACCAAT 1108
||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 682 AspPheLysLeuTyrAlaSerAlaPheSerGluArgTyrLeuGlyLeuHisGlyLeuAsp 701
QY 1109 GAACAGGGTATATCTAGGACTCTGGCCATGACGACGAGAAAGTCCCTCTCGAACCA 1168
||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 702 AsnArgAlaTyrGluMetAlaLysValAlaHisArgValSerAlaLeu-----GluGly 719
QY 1169 AATCGTTTACTGCTCTTACATGGTTTCTGTGATGAGAAATGTCATTTTGCATACACAGT 1228
||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 720 GlnGlnPheLeuValIleHisAlaThrAlaAspGluLysIleHisPheGlnHisThrAla 739
QY 1229 ATATTACTAGTGTATTTAGTGGGCTCGAAAGCCATATGATTTTACAGATCTATCTCTCAG 1288
||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 740 GluLeuIleThrGlnLeuIleLysGlyLysAlaAsnTyrSerLeuGlnIleTyrProAsp 759
QY 1289 GAGAGACACAGCATAAAGAGTTCTCTGAATCGGGAGAACATATTGAACTGCATCTTTTGCAC 1348
||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 760 GluSerHisTyrPheSerSerAlaAlaLeuGlnGlnHisLeuHisArgSerIleLeuGly 779
```

QY 1349 TACCTTCAGAA 1360
Db 780 PhePheValGlu 783
RESULT 14
168600
dipeptidyl aminopeptidase like protein - human
C;Species: Homo sapiens (man)
C;Date: 06-Sep-1996 #sequence_revision 06-Sep-1996 #text_change 21-Jul-2000
R;Yokotani, N.; Doi, K.; Wenthold, R.J.; Wada, K.
Hum Mol Genet. 2, 1037-1039, 1993
A;Title: Non-conservation of a catalytic residue in a dipeptidyl aminopeptidase IV-related
A;Reference number: 154331; MUID:93372805; PMID:810397
A;Accession: I68600
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-803 <RES>
A;Cross-references: UNIPARC:UPI000016A7F9; GB:M96860; NID:G306707; PIDN:AAA35761.1; PID:
C;Superfamily: dipeptidyl-peptidase IV

Alignment Scores:
Pred. No.: 4,45e-10 Length: 803
Score: 216.00 Matches: 111
Percent Similarity: 32.7% Conservative: 79
Best Local Similarity: 19.1% Mismatches: 179
Query Match: 7.3% Indels: 212
DB: 2 Gaps: 23

US-10-825-632-6 (1-1669) x I68600 (1-803)

QY 5 GGTACAGCAATCTTAAGTCACTTTAAGTGTGAGAAATGATGCTGAGGA 64
Db 286 GlySerGluAsnProSerLeuHis-----ValIleGlyLeuAsnGly 301
QY 65 AGGATCATAGTCTCATAGTAAGCACTAATTCACCTTTTGAGATTCTATTGGAAGGA 124
Db 302 ProThrHisAspLeu-----GluMetMetProProAspAspProArgMetArgGlu 318
QY 125 GTTGAATATATGCGAGAGCTGGATGACTCTCGAGGAAATATGCT-----TGFTCC 178
Db 319 Tyr---TyrIleThrMetValLysTrpAlaThrSerThrLysValAlaValThrTrp--- 336
QY 179 ATCTCATAGATCGCTCCAGATCCGCTACAGATAGTGTGATCTACCTGATTTATTT 238
Db 337 -----LeuAsnArgAlaGlnAsn-----ValSer----- 344
QY 239 ATCCAGTAGAATGATGTTATGGAAGGAGAGACTCATTGAGTCACTGCTGATTC 298
Db 345 -----IleLeuThrLeuCysAspAla 351
QY 299 GTGACGCCACTAATATTC-----TATGAAGAAACACAGACATCTGGATAAATATCCAT 352
Db 352 ThrThrGlyValCysThrLysLysHisGluAspGluSerGluAlaTrpLeuHis----- 369
QY 353 GACATCTTTCATGTTTCCCAAGTCACGAAGAGGAAAT----- 394
Db 370 -----ArgGlnAsnGluGluProValPheSerLysAspGlyArg 382
QY 395 GAGTTTATTTTGCTCTGATGCAAAACAGGTTTCCGT-----CATTTATACAAATATACA 451
Db 383 LysPhePheIleArgAlaIleProGlnGlyGlyArgGlyLysPheThrHisIleThr 402
QY 452 TCTATTTTAAAGGAAGCAATATATAACGATCGAGTGGTGGGCTGCTGCTCCCAAGTAT 511
Db 403 -----ValSerSerSerGlnProAsnSerSerAsnAspAsnIleGln----- 416
QY 512 TTCAGTGTCTCATCAAGAGGAGATAGCAATTTACAGTGGTGAATGGAGTCTTGGC 571
Db 417 -----SerIleThrSerGlyAspTrpAspValThr----- 426
QY 572 CGGATGGATCTAATATCCAAAGTGTGAGTCAAGGAGCTGGTATATTGGAAGGCACC 631
Db 573 ----- 631

Db 427 -----LysIleLeuAlaTyrAspGluLysGlyAsnLysIleTyrPheLeuSerThr 443
QY 632 AAAGACTCCCTTTAGAGCATCACCTGTACGTAGCTAGCTAGCTAAATCTCGAGAGGTG 691
Db 444 GluAspLeuProArgArgGlnLeuTyrSerAlaAsn----- 456
QY 692 ACAAGGCTGACGACCGTGGCTACTCACAT-----TCTTCTGCATCAGTCAGCAC 742
Db 457 -----ThrGluGlyAsnPheAsnArgGlnCysLeuSerCysAspLeuValGluAsn 473
QY 743 TGTGACTCTTTTATAGTAAGTATAGTAACAGAGAATCCACACTGTGTGCTCCTTAC 802
Db 474 CysThrTyrPheSerAlaSerPheSerHisSerMetAsp-----PhePheLeuLeu 490
QY 803 AAGCTATCAAGTCT-----GAAGATGACCCCACTTGCCTGCAAAACAAG 844
Db 491 LysCysGluGlyProGlyValProMetValThrValHisAsnThrThrAspLysLysLys 510
QY 845 GAATTT-----TGGGCCACCATTTTGGATTTCAGCAGGTCT 880
Db 511 MetPheAspLeuGluThrAsnGluHisValLysLysAlaIleAsnAspArgGlnMetPro 530
QY 881 CTTCCTGACTATCTCTCCAGAAATTTCTCTTTGAAAGTACTTCTGGAATTTACATTG 940
Db 531 LysValGluTyrArgAspIleGluIle-----AspAspTyrAsnLeu 544
QY 941 TATGGGATGCTCTACAGGCTCATGATCTACAGCCTGGAAATATCTCTACTGTGCTG 1000
Db 545 ProMetGlnIleLeuLysProAlaThrPheThrAspThrThrHisTyrProLeuLeuLeu 564
QY 1001 TTCATATATGCTGCTCT----- 1018
Db 565 ValValAspGlyThrProGlySerGlnSerValAlaGluLysPheGluValSerTrpGlu 584
QY 1018 ----- 1018
Db 585 ThrValMetValSerSerHisGlyAlaValValLysCysAspGlyArgGlySerGly 604
QY 1018 ----- 1018
Db 605 PheGlnGlyThrLysLeuLeuHisGluValAlaArgArgLeuGlyLeuGluGluLys 624
QY 1019 -----CAGTT 1024
Db 625 AspGlnMetGluAlaValArgThrMetMetLeuLysGluGlnTyrIleAspArgThrArgVal 644
QY 1025 GCTATTGCTGG----- 1036
Db 645 AlaValPheGlyLysAspTyrGlyTyrLeuSerThrTyrIleLeuProAlaLysGly 664
QY 1037 -----GCCCCAGTCACTCTGTGGATC 1057
Db 665 GluAsnGlnGlyGlnThrPheThrCysGlySerAlaLeuSerProIleThrAspPheLys 684
QY 1058 TTCTATGATACAGGATACACGGAACGTTATATGGTCACTCCCTGACCAAGTAACAGGC 1117
Db 685 LeuTyrAlaSerAlaPheSerGluArgTyrLeuGlyLeuHisGlyLeuAspAsnArgAla 704
QY 1118 TATTACTTAGATCTGGCCATGCAAGCAAGAAAGTCTCCCTCTGACCAAAATCGTTA 1177
Db 705 TyrGluMetThrLysValAlaHisArgValSerAlaLeu-----GluGlnGlnPhe 722
QY 1178 CTGCTCTTACATGTTTCTCTGGATGAGAATGTCATTTTGCATACCATGATATATCTG 1237
Db 723 LeuIleIleHisProThrAlaAspGluLysIleHisPheGlnHisThrAlaGluLeuIle 742
QY 1238 AGTTTTATGAGGCTGGAAAGCCCATATGATTTTACAGATCTATCTCTCAGAGACAC 1297
Db 743 ThrGlnLeuIleArgGlyLysAlaAsnTyrSerLeuGlnIleTyrProAspGluSerHis 762
QY 1298 ACATAAGAGTTCCTGAATCGGAGAACATATGAACATGCTATCTTTTGCACCTTCAA 1357
Db 763 TyrPheThrSerSerSerLeuLysGlnHisLeuTyrArgSerIleIleAsnPheVal 782

```
QY 1358 GAA 1360
Db 783 Glu 783
RESULT 15
I54331
dipeptidyl aminopeptidase like protein - human
C:Species: Homo sapiens (man)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
C:Accession: I54331
R:Yokotani, N.; Doi, K.; Wenthold, R.J.; Wada, K.
Hum. Mol. Genet. 2, 1037-1039, 1993
A:Title: Non-conservation of a catalytic residue in a dipeptidyl aminopeptidase IV-related
A:Reference number: I54331; MUID:9372805; PMID:8103397
A:Accession: I54331
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-865 <RES>
A:Cross-references: UNIPROT:P42658; UNIPARC:UPI0000047378; GB:M96859; NID:g306705; PIDN:
C:Superfamily: dipeptidyl-peptidase IV

Alignment Scores:
Pred. No.: 4.54e-10 Length: 865
Score: 216.00 Matches: 111
Percent Similarity: 32.7% Conservative: 79
Best Local Similarity: 19.1% Mismatches: 179
Query Match: 7.3% Indels: 212
DB: 2 Gaps: 23

US-10-825-632-6 (1-1669) x I54331 (1-865)
QY 5 GGTACAGCAAACTCTAAAGTCACTTTTAAGATGTGAGAAATATGATGTCGAAGGA 64
Db 348 GlySerGluAsnProSerIleSerLeuHis-----ValIleGlyLeuAsnGly 363
QY 65 AGCATATAGATCATAGATAGGAAGTAACTTTCAGCTTTTGAGATCTTATTGAGGA 124
Db 364 ProThrHisAspLeu-----GluMetProProAspAspProArgMetArgGlu 380
QY 125 GTTGAATATATGCCAGAGCTGGAGTCTCTCGAGGAAATATGCT-----TGGTCC 178
Db 381 Tyr---TyrIleThrMetValIleStrpAlaThrSerThrIleValAlaValThrTrp--- 398
QY 179 ATCTACTAGATCGCTCCAGACTCGCTCAGATAGTGTGATCTCAGCTGAATATTT 238
Db 399 -----LeuAsnArgAlaGlnAsn-----ValSer----- 406
QY 239 ATCCAGTAGAGATGATGTTATGGAAGGAGAGACTCATTTGAGTCAGTCTGATCT 298
Db 407 -----ArgGlnAsnGluGluProValPheSerLysAspGlyArg 413
QY 299 GTGACGCCCACTAATTATC-----TATGAAGAAACAACAGACATCTGGATAATATCCAT 352
Db 414 ThrThrGlyValCysThrLysIleGluAspGluSerGluAlaTrpLeuHis----- 431
QY 353 GACATCTTTTCATGTTTTTCCCAAGTCCAGAGAGGAAAT----- 394
Db 432 -----ArgGlnAsnGluGluProValPheSerLysAspGlyArg 444
QY 395 GAGTTTATTTTTCCTCTGAATGCAAAACAGGTTTCCGT---CATTTATACAAATTACA 451
Db 445 LysPhePheIleArgAlaIleProGlnGlyIleArgGlyLysPheThrHisIleThr 464
QY 452 TCTATTATAAGGAAAGCAATATATAACGATCCAGTGGTGGTGGTCTCTCCAAGTAT 511
Db 465 -----ValSerSerSerGlnProAsnSerSerAsnAspAsnIleGln----- 478
QY 512 TTCAAGTGTCTATCAAGAGAGATAGCAATTACCAGTGGTGAATGGAGGTTCTTGGC 571
Db 479 -----SerIleThrSerGlyAspTrpAspValThr--- 488
QY 572 CGGCATGATCTAATATCCAAAGTTGATGAAGTCAGAGGCTGGTATATTTGAAGGCACC 631

489 -----LysIleLeuAlaTyrAspGlyLysGlyAsnLysIleTyrPheLeuSerThr 505
632 AAAGACTCCCTTTAGAGCATCACCTGTAGTACGTAGTACGTAAATCCTGAGAGGTG 691
506 GluAspLeuProArgArgGlnLeuTyrSerAlaAsn----- 518
692 ACAAGGCTGACTGACCGTGTACTACAT-----TCTTGTGTCATCAGTCAGCAC 742
519 -----ThrGluGlyAsnPheAsnArgGlnCysLeuSerCysAspLeuValGluAsn 535
743 TGTGACTCTTTTATAAGTAAGTATAGTAACACAGAAATCCACACTGTGTGTCCTTTAC 802
536 CysThrTyrPheSerAlaSerPheSerHisSerMetAsp-----PhePheLeuLeu 552
803 AAGCTTATCAAGTCTCT-----GAAGATGACCCCACTTGCAGAAACAAG 844
553 LysCysGluGlyProGlyValProMetValThrValHisAsnThrAspLysLysLys 572
845 GAATTT-----TGGGCCACCATTTTGGATTCACAGGTCTCT 880
573 MetPheAspLeuGluThrAsnGluHisValLysLysAlaIleAsnAspArgGlnMetPro 592
881 CTTCTCTGACTATCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGATTACATTG 940
593 LysValGluIleArgAspIleGluIle-----AspAspTyrAsnLeu 606
941 TATGGGATGCTCTACAGCTCATGATCTACAGCTGGAAAGAAATATCTCTACTGTGCTG 1000
607 ProMetGlnIleLeuLysProAlaThrPheThrAspThrThrHisTyrProLeuLeu 626
1001 TTCATATGCTGGTCTCT----- 1018
627 ValValAspGlyThrProGlySerGlnSerValAlaGluLysPheGluValSerTrpGlu 646
1018 ----- 1018
647 ThrValMetValSerSerHisGlyAlaValValLysCysAspGlyArgGlySerGly 666
1018 ----- 1018
667 PheGlnGlyThrLysLeuLeuHisGluValArgArgArgLeuGlyLeuLeuGluLys 686
1019 -----CAGGTT 1024
687 AspGlnMetGluAlaValArgThrMetLeuLysGluGlnTyrIleAspArgThrArgVal 706
1025 GCTATTGCTGGG----- 1036
707 AlaValPheGlyLysAspTyrGlyGlyTyrLeuSerThrTyrIleLeuProAlaLysGly 726
1037 -----GCCCAAGTCACTCTGTGATC 1057
727 GluAsnGlnGlyGlnThrPheThrCysGlySerAlaLeuSerProIleThrAspPheLys 746
1058 TTCTATGATACAGGATACACGGAACGTTATATGGTCCACCTGACCCAGAAATGAACAGGC 1117
747 LeuTyrAlaSerAlaPheSerGluArgTyrLeuGlyLeuHisGlyLeuAspAsnArgAla 766
1118 TATTACTTAGGATCTGTGGCCATCGACGAGAAAGTTCCCTCTCTGCAACCAATCTTTA 1177
767 TyrGluMetThrLysValAlaHisArgValSerAlaLeu-----GluGluGlnGlnPhe 784
1178 CTGCTCTTACATGGTTCTCTGGTGGAGATGTCTCTTTTGCATACATCCAGTATATTACTG 1237
785 LeuIleIleHisProThrAlaAspGluLysIleHisPheGlnHisThrAlaGluLeuIle 804
1238 AGTTTATTAGTGGGCTGGAAAGCCATATGATTTACAGATCTATCTCTCAGGAGAGACAC 1297
805 ThrGlnLeuIleArgGlyLysAlaAsnTyrSerLeuGlnIleTyrProAspGluSerHis 824
1298 AGCATAGAGTCTCTGAATCGGAGAACATTATTAAGTCTGTCATCTTTTGCACCTACCTCAA 1357
-----
```

Db 825 TyrPheThrSerSerLeuLysGlnHisLeuTyrArgSerIleIleAsnPheVal 844

Qy 1358 GAA 1360

|||

Db 845 Glu 845

Search completed: May 2, 2006, 01:37:46
Job time : 89.6336 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: May 2, 2006, 00:42:23 ; Search time 76.878 Seconds
(without alignments)
3063.366 Million cell updates/sec

Title: US-10-825-632-6
Perfect score: 2957
Sequence: 1 aacagtcagcaaatccta.....aaaaaaaaaaaaaaaaaaaaa 1669

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 4332886

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEl=frame+ n2p.model -DEV=xlbp
-Q=/abs/ABSSWEB.spool/US10825632/runat_01052006_105946_3226/app_query.fasta-1
-DB=UniProt -QMT=faetan -SUFFIX=n2p.rup -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blom62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFWT=ptp -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abs02p
-USER=US10825632.OCGN 1_1_694 @runat_01052006_105946_3226 -NCPU=6 -ICPU=3
-NO WMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : UniProt_05.80.*

1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2422	81.9	898	1 DPP8 HUMAN	Q6v1x1 homo sapien
2	2326	78.7	892	1 DPP8_MOUSE	Q80ya7 mus musculu
3	1494.5	50.5	863	1 DPP9 HUMAN	Q86ti2 homo sapien
4	1488.5	50.3	862	1 DPP9_MOUSE	Q8bvg4 mus musculu
5	1462	49.4	923	2 Q4SBM6 TETNG	Q4sbm6 tetraodon n
6	1389	47.0	847	2 Q6R222_XENLA	Q6gr22 xenopus lae
7	1187.5	40.2	508	2 Q75273_HUMAN	Q75273 homo sapien
8	894	30.2	621	2 Q7PTT8_ANOGA	Q7ptt8 anopheles g
9	889	30.1	886	2 Q7QBK1_ANOGA	Q7qbk1 anopheles g
10	867.5	29.3	740	2 Q5TTK8_ANOGA	Q5ttk8 anopheles g
11	841.5	28.5	1053	2 Q9VC20_DROME	Q9vc20 drosophila
12	841.5	28.5	1113	2 Q9VCL9_DROME	Q9vc19 drosophila
13	718.5	24.3	557	2 Q5TXJ2_ANOGA	Q5txj2 anopheles g
14	613.5	20.7	803	2 Q54U01_DICDI	Q54u01 dictyosteli
15	462	15.6	927	2 Q965K3_CABEL	Q965k3 caenorhabdi
16	462	15.6	931	2 Q44987_CABEL	Q44987 caenorhabdi

17	449	15.2	552	2	Q8GUJ7_ARATH	Q8guj7 arabidopsis
18	449	15.2	746	2	Q9FNF6_ARATH	Q9fnf6 arabidopsis
19	444	15.0	596	2	Q6K880_ORYZA	Q6k880 oryza sativ
20	429.5	14.5	738	2	Q9A6E0_CAUCR	Q9a6e0 caulobacter
21	415.5	14.1	763	2	Q8EAB7_SHEON	Q8eab7 shewanella
22	413	14.0	895	2	Q61CU7_CABBR	Q61cu7 caenorhabdi
23	391	13.2	743	2	Q5QX36_IDILO	Q5qx36 idiomarina
24	386	13.1	238	2	Q5TYG2_ANOGA	Q5tyg2 anopheles g
25	383	13.0	720	2	Q5NMW8_ZYMMO	Q5nmw8 zymomonas m
26	379.5	12.8	745	2	Q6F317_P9BED	Q6f317 pseudomonas
27	369.5	12.5	736	2	Q8A028_BACTN	Q8a028 bacteroides
28	359.5	12.2	736	2	Q5LGU5_BACPH	Q5lgu5 bacteroides
29	359.5	12.2	766	2	Q4TNP1_9SPHN	Q4tnp1 erythrobact
30	356.5	12.1	736	2	Q64XP9_BACFR	Q64xp9 bacteroides
31	355.5	12.0	751	2	Q4UFD3_XANCP	Q4upd3 xanthomonas
32	355.5	12.0	751	2	Q8P3V8_XANCP	Q8p3v8 xanthomonas
33	355	12.0	741	2	P95782_XANNA	P95782 xanthomonas
34	351	11.9	745	2	Q5H5W8_XANOR	Q5h5w8 xanthomonas
35	346	11.7	757	2	Q8PFD7_XANAC	Q8pfd7 xanthomonas
36	322	10.9	769	2	Q5QVY7_IDILO	Q5qvyy7 idiomarina
37	320.5	10.8	850	2	Q6H9E3_9TRYP	Q6h9e3 trypanosoma
38	318	10.8	947	2	Q4WX13_ASPFU	Q4wx13 aspergillus
39	312.5	10.6	874	2	Q7SI80_EMENI	Q7si80 emericella
40	312.5	10.6	880	2	Q5B934_EMENI	Q5b934 aspergillus
41	312	10.6	901	2	Q96VT7_ASPNG	Q96vt7 aspergillus
42	309.5	10.5	771	2	Q42812_ASPNG	Q42812 aspergillus
43	308.5	10.4	765	2	Q4WPH9_ASPFU	Q4wph9 aspergillus
44	308.5	10.4	765	2	O14425_ASPFU	O14425 aspergillus
45	306.5	10.4	778	2	Q51ZP7_MAGGR	Q51zp7 magnaporthe

ALIGNMENTS

RESULT 1				
ID	DPP8 HUMAN	STANDARD;	PRT;	898 AA.
AC	Q6VLX1; Q7Z4C8; Q7Z4D3; Q7Z4E1; Q81WG7; Q8NEM5; Q96JX1; Q9HEM2;			
AC	Q9HEM3; Q9HEM4; Q9HEM5; Q9NXP4;			
DT	13-SEP-2005 (Rel. 48, Created)			
DT	13-SEP-2005 (Rel. 48, Last sequence update)			
DT	13-SEP-2005 (Rel. 48, Last annotation update)			
DE	Dipeptidyl peptidase 8 (EC 3.4.14.5) (Dipeptidyl peptidase VIII) (DP8)			
DE	(Prolyl dipeptidase DPP8) (Dipeptidyl peptidase IV-related protein 1) (DPP8-1).			
GN	Name=DPP8; Synonyms=DPP81; ORFNames=MSTP097, MSTP135, MSTP141;			
OS	Homo sapiens (Human)			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;			
OC	Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 3), NUCLEOTIDE SEQUENCE [MRNA] OF 334-898 (ISOFORM 4), NUCLEOTIDE SEQUENCE [MRNA] OF 540-898 (ISOFORM 5), NUCLEOTIDE SEQUENCE [MRNA] OF 260-792 (ISOFORM 6), FUNCTION, CATALYTIC ACTIVITY, ENZYME REGULATION, TISSUE SPECIFICITY, INDUCTION, AND SUBCELLULAR LOCATION.			
RC	TISSUE=Placenta;			
RX	MEDLINE=20467194; PubMed=11012666;			
RA	Abbott C.A., Yu D.M.T., Woollatt E., Sutherland G.R., McCaughan G.W.,			
RA	Correll M.D.;			
RT	"Cloning, expression and chromosomal localization of a novel human dipeptidyl peptidase (DPP) IV homolog, DPP8.,"			
RL	Eur. J. Biochem. 267:6140-6150(2000).			
RN	[2]			
RP	NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 3), CATALYTIC ACTIVITY, ENZYME REGULATION, BIOPHYSICOCHEMICAL PROPERTIES, SUBCELLULAR LOCATION, AND TISSUE SPECIFICITY.			
RC	TISSUE=Testis;			
RX	PubMed=12662155; DOI=10.1042/BJ20021914;			
RA	Qi S.Y., Riviere P.J., Trojnar J., Junien J.-L., Akinesanya K.O.;			
RT	"Cloning and characterization of dipeptidyl peptidase 10, a new member of an emerging subgroup of serine proteases.,"			
RL	Biochem. J. 373:179-189(2003).			

[3]
RN NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 1).
RP TISSUE=Testis;
RA Sha J.H., Zhou Z.M., Li J.M.;
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
RN [4]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 4), AND NUCLEOTIDE
RP SEQUENCE [LARGE SCALE MRNA] OF 211-898 (ISOFORM 2).
RC TISSUE=Hepatoma, and Placenta;
RX PubMed=14702039; DOI=10.1038/ng1285;
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA Sekine M., Ohtsuka M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA Yamamoto J.-I., Saito K., Yasui T., Isono Y., Nakamura Y.,
RA Nagahashi K., Murakami K., Kawada T., Iwayanagi T., Wagatsuma M.,
RA Shiratori A., Sudo H., Hosioki T., Kaku Y., Kodaira H., Kikkawa E.,
RA Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E.,
RA Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa K.,
RA Yamazaki M., Ninomiya K., Ishibashi T., Yamashita H., Murakawa K.,
RA Fujimori K., Tanai H., Kimata M., Watanabe M., Hirakawa S., Chiba Y.,
RA Ishida S., Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T.,
RA Kusano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,
RA Nomura Y., Togihara S., Komai F., Hara R., Takeuchi K., Arita M.,
RA Imoe N., Mutsaers H., Uehira H., Oshima A., Sasaki N., Aotsuka S.,
RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohara N., Sano S.,
RA Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Suzuki H.,
RA Hishigaki H., Watanabe K., Kumagai A., Itakura S., Fukuzumi B.,
RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Omori Y., Okamoto S.,
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikemura Y., Okamoto S.,
RA Okitani K., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
RT "Complete sequencing and characterization of 21,243 full-length human
RT cDNAs.";
RL Nat. Genet. 36:40-45(2004).
[5]
RN [6]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORMS 1 AND 2).
RC TISSUE=Testis;
RX MEDLINE=2388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.P., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Raba S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Rahay J., Helton E., Kettunen M., Madan A., Rodriguez S., Sanchez A.,
RA Whitling M., Madan A., Young A.C., Green E.D., Dickson M.C.,
RA Blakeley R.W., Touchman J.W., Schmutz J., Myers R.M.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Skalek U., Smallus D.B.,
RA Butterfield J.S.N., Krzywicki M.I., Skalek U., Smallus D.B.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[6]
RN [7]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 561-898.
RC TISSUE=Aorta;
RA Zhao B., Xu H.S., Tong Y.K., Sheng H., Qin B.M., Liu Y.Q., Liu B.,
RA Wang X.Y., Zhang Q., Song L., Gao Y., Zhang C.L., Ye J., Ji X.J.,
RA Liu B.H., Lu H., Chen J.Z., Cai M.Q., Zheng W.Y., Teng C.Y., Liu Q.,
RA Yu L.T., Lin J., Gong Q., Zhang A.M., Gao R.L., Hui R.T.;
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
[7]
RN [8]
RP MUTAGENESIS OF GLU-275; SER-755; ASP-833 AND HIS-865, CATALYTIC
RP ACTIVITY, BIOPHYSICOCHEMICAL PROPERTIES, AND SUBCELLULAR LOCATION.
RX PubMed=12534281; DOI=10.1021/bi026946g;
RA Ajami K., Abbott C.A., Obradovic M., Gysbers V., Kaehne T.,
RA McCaughan G.W., Gorrell M.D.;
RT "Structural requirements for catalysis, expression, and dimerization
RT in the CD26/DPPIV gene family.";
RL Biochemistry 42:694-701(2003).
RN [9]
RP BIOPHYSICOCHEMICAL PROPERTIES.
RX PubMed=15039077; DOI=10.1016/j.pep.2003.12.019;
RA Chen Y.-S., Chien C.-H., Goparaju C.M., Hsu J.T.-A., Liang P.-H.,
RA Chen X.;
RT "Purification and characterization of human prolyl dipeptidase DPP8 in
RT Sf9 insect cells.";
RL Protein Expr. Purif. 35:142-146(2004).
RN [9]
RP ENZYME REGULATION.
RX PubMed=15664838; DOI=10.1016/j.bmcl.2004.11.023;
RA Jiaang W.-P., Chen Y.-S., Hsu T., Wu S.-H., Chien C.-H., Chang C.-N.,
RA Chang S.-P., Lee S.-J., Chen X.;
RT "Novel isoindoline compounds for potent and selective inhibition of
RT prolyl dipeptidase DPP8.";
RL Bioorg. Med. Chem. Lett. 15:687-691(2005).
CC -!- FUNCTION: Dipeptidyl peptidase that cleaves off N-terminal
CC dipeptides from proteins having a Pro or Ala residue at position
CC 2. May play a role in -cell activation and immune function.
CC -!- CATALYTIC ACTIVITY: Release of an N-terminal dipeptide, Xaa-Yaa-|-
CC Zaa-, from a polypeptide, preferentially when Yaa is Pro, provided
CC Zaa is neither Pro nor hydroxyproline.
CC -!- ENZYME REGULATION: Inhibited by zinc. Inhibited by the serine
CC proteinase inhibitor 4-(2-aminoethyl)benzenesulphonyl fluoride
CC (ABEPSF), and by di-isopropylfluorophosphate. Specifically inhibited
CC by isoindoline derivatives.
CC -!- BIOPHYSICOCHEMICAL PROPERTIES:
CC Kinetic parameters:
CC KM=208 uM for Ala-Pro-AMC;
CC KM=130 uM for Ala-Pro-AFC;
CC KM=120 uM for H-Ala-Pro-pNa;
CC KM=1420 uM for H-Ala-Ala-pNa;
CC KM=310 uM for H-Arg-Pro-pNa;
CC KM=2050 uM for H-Asp-Pro-pNa;
CC KM=480 uM for H-Gly-Pro-pNa;
CC pH dependence:
CC Optimum pH is 7.4-8.5. Little activity below pH 6.5;
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=6;
CC Name=1;
CC IsoId=06VIX1-1; Sequence=Displayed;
CC Name=2;
CC IsoId=06VIX1-2; Sequence=VSP_013864;
CC Name=3;
CC IsoId=06VIX1-3; Sequence=VSP_013860;
CC Name=4;
CC IsoId=06VIX1-4; Sequence=VSP_013860, VSP_013862;
CC Name=5;
CC IsoId=06VIX1-5; Sequence=VSP_013863;
CC Name=6;
CC IsoId=06VIX1-6; Sequence=VSP_013861;
CC -!- TISSUE SPECIFICITY: Ubiquitously expressed, with highest levels in
CC testis, placenta, prostate, muscle and brain.
CC -!- INDUCTION: In activated T-cells.
CC -!- SIMILARITY: Belongs to the peptidase S9B family. DPPIV subfamily.
CC -!- CAUTION: Ref.4 (BAA91059) sequence differs from that shown due to
CC a frameshift in position 486.
CC -!- CAUTION: Ref.6 (AAQ13650 and AAQ13623) sequences differ from that
CC shown due to several frameshifts.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

AC Q80YA7; Q9D4G6;
DT 13-SEP-2005 (Rel. 48, Created)
DT 13-SEP-2005 (Rel. 48, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Dipeptidyl peptidase 8 (EC 3.4.14.5) (Dipeptidyl peptidase VIII)
DE (D8)
GN Name=Dpp8;
OS Mus musculus (Mouse)
OC Eukaryota; Metazoa;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Xiphiidae; Muridae; Murinae; Mus.
OX [1]
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
RA Nakai Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nakaide I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schrim L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais K., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertes G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramchandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B., Ringwald M.,
RA Sandelin A., Schneider C., Semp C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Wahlstedt S., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wyshak B., Zeng L., Yanagisawa M., Yang L., Yang L.,
RA Yuan Z., Zeng L., Zeng L., Zeng L., Zeng L., Zeng L., Zeng L.,
RA Hirose K., Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sakaki D., Shibata K., Shinagawa A.,
RA Tsunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573 (2002).
RN [2]
RN NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RP STRAIN=C57BL/6; TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klauener R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.P., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heish F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mallya S.J.,
RA Besak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Blakesley R.W., Tomich J.W., Green E.D., Dickinson M.C.,
RA Roderick T., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywicki M.I., Skalska U., Smallus D.E.,
RA Schneider A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [3]
RN PHOSPHORYLATION SITE TYR-325.
RP PubMed=14729942; DOI=10.1074/mcp.D300003-MCP200;
RX Shu H., Chen S., Bi Q., Mumby M., Brekken D.L.;
RT "Identification of phosphoproteins and their phosphorylation sites in
RT the WEHI-231 B lymphoma cell line."
RL Mol. Cell. Proteomics 3:279-286 (2004).
CC -!- FUNCTION: Dipeptidyl peptidase that cleaves off N-terminal
CC dipeptides from proteins having a Pro or Ala residue at position
CC 2. May play a role in T-cell activation and immune function (By
CC similarity).
CC -!- CATALYTIC ACTIVITY: Release of an N-terminal dipeptide, Xaa-Yaa-|-
CC zaa-, from a polypeptide, preferentially when Yaa is Pro, provided
CC zaa is neither Pro nor hydroxyproline.
CC -!- ENZYME REGULATION: Inhibited by zinc. Inhibited by the serine
CC proteinase inhibitor 4-(2-aminoethyl)benzenesulfonyl fluoride
CC (AEBSF), and by di-isopropyl fluorophosphate. Specifically inhibited
CC by isocitidine derivatives (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the peptidase S9B family. DPPIV subfamily.
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; AK016546; BAB30295.2; -; mRNA.
CC EMBL; BC043124; AAH43124.1; -; mRNA.
CC EMBL; BC059222; AAH59222.1; -; mRNA.
CC KEROFS; S09.018; -; -;
CC Ensembl; ENSMUSG00000032393; Mus musculus.
CC MG1; MG1:1921638; DPP8.
CC InterPro; IPR001375; Peptidase_S9.
CC InterPro; IPR002469; Peptidase_S9B.
CC InterPro; IPR000379; Ser_estrs.
CC Pfam; PF00930; DPPIV_N; 1.
CC Pfam; PF00326; Peptidase_S9; 1.
CC KW Amino peptidase; Hydrolase; Phosphorylation; Protease; Serine protease.
FT ACT SITE 749 749 Charge relay system (By similarity).
FT ACT SITE 827 827 Charge relay system (By similarity).
FT ACT SITE 859 859 Charge relay system (By similarity).
FT MOD RES 325 325 Phosphotyrosine.
FT CONFLICT 87 87 G -> R (in Ref. 1).
SQ SEQUENCE 892 AA; 102186 MW; 59081CD9792E03ED CRC64;
Alignment Scores:
Pred. No.: 3,18e-187 Length: 892
Score: 2326.00 Matches: 443
Percent Similarity: 80.2% Conservative: 10
Best Local Similarity: 78.4% Mismatches: 12
Query Match: 78.7% Indels: 100
DB: Gaps: 1
US-10-825-632-6 (1-1669) x DPP8_MOUSE (1-892)
QY 2 ACAGGTACGACCAATCCTTAAGTCACTTTTAAAGTGTGTCAGAAATAATGATGCTGAA 61
Db 328 ThrGlyThrAlaAsnProlysalThrPhelyMetSerGluIleValIaValaPala 347
QY 62 GGAAGATCATGATGTCATAGATGAAGAACTAATCAACCTTTTGAGATTCATTGAA 121
Db 348 GlyGlyIleIleAspValIleAspGluLeuValGlnProPheGluIleLeuPheGlu 367
QY 122 GGAGTTGAATATATTGCCAGAGCTGGAGTGGACTCTCTGAGGAAATATGCTTGGTCCATC 181
Db 368 GlyValGlyIleIleAlaArgAlaGlyIleThrProGluGlyIleHisAlaIleTrpSerIle 387
QY 182 CTACTAGATCGCTCCAGAGCTCGCTCAGATAGTGTGATCTCCACCTGAATATTATTC 241
Db 388 LeuLeuAspArgSerGlnThrHisLeuGlnIleValLeuIleSerProGluLeuPheIle 407
QY 242 CCAGTGAAGATGATGTTATGAAAGGCAGAGACTCATTGAGTCAGTCCCTGATTCGTG 301
Db 408 ProValGluAspAlaMetAspArgGlnArgLeuIleGluSerValProAspSerVal 427
QY 302 ACGCCACTAATTATCTATGAAGAAACACAGACATCTGGATAATAATATCCATCTTT 361

Db 428 ThrProLeuIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePhe 447
QY 362 CATGTTTTCCTCCCAAGTACAGAGAGAAATTTAGTTTATTTTGGCTCTCTGAATGCATA 421
Db 448 HisValPheProGlnThrHisGluAspGluIleGluPheIlePheAlaSerGluCysLys 467
QY 422 ACAGTTTCCTCATTTATACAAAATTACATCTATTTAAAGAAAGCAATATAACGA 481
Db 468 ThrGlyPheArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArg 487
QY 482 TCCAGTGGTGGCTGCTCTCCAGTGTCTCAAGTGTCTCTATCAAGAGGAGATAGCA 541
Db 488 SerSerGlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleThr 507
QY 542 ATTACAGTGGTGAATGGGAAGTTCTTGGCCGGCATGATCAATATCAAGTTGATGAA 601
Db 508 IleThrSerGlyGluTrpGluValLeuGlyArgHisGlySerAsnIleTrpValAspGlu 527
QY 602 GTGAGAGGCTGTATATTTGAGGACCAAGACTCCCTTTAGAGCATCACTGTAC 661
Db 528 AlaArgLysLeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLysLeuTyr 547
QY 662 GTAGTCAGTTACGTAATCTCGAGAGGTGACAAAGCTGACTGACCGTGGCTACTACAT 721
Db 548 ValThrSerTyrAlaAsnProGlyGluValValArgLeuThrAspArgGlyTyrSerHis 567
QY 722 TCTGTGTCATAGTCAGTCAGTGTGCTTTTATTAAGTAAATAGTAAACGAGAAAT 781
Db 568 SerCysLysSerArgHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsn 587
QY 782 CCACACTGTGTGCTTACAGCTATCAGTCTGAGATGACCCAACTGCGAAACA 841
Db 588 ProHisCysValSerLeuTyrLysLeuSerProGluAspAspProValHisLysThr 607
QY 842 AAGGAATTTGGGCCACCATTTGGATTGAGTTCAGCAGGTCTCTCTCTGACTACTCTCCA 901
Db 608 LysGluPheTrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProPro 627
QY 902 GAAATTTCTTTGAAAGTACTACTGATTTACATTTGATGGATGCTCTACAGCCT 961
Db 628 GluIlePheSerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysPro 647
QY 962 CATGATCTACAGCTCGAAGAAATATCTACTGTGCTGTTCATATATGGTGGCTCAG 1021
Db 648 HisAspLeuGlnProGlyLysTyrProThrValLeuPheIleTyrGlyGlyProGln 667
QY 1021 ----- 1021
Db 668 ValGlnLeuValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAla 687
QY 1021 ----- 1021
Db 688 SerLeuGlyTyrValValValValIleAspAsnArgGlySerCysHisArgGlyLeuLys 707
QY 1021 ----- 1021
Db 708 PheGluGlyAlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGly 727
QY 1021 ----- 1021
Db 728 LeuGlnTyrLeuAlaSerGlnTyrAspPheIleAspLeuAspArgValGlyIleHisGly 747
QY 1021 ----- 1021
Db 748 TrpSerTyrGlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArg 767
QY 1022 GTTGCTATTGCTGGGCCCCAGTCACTCTGTGGATCTCTATGATACAGGATACACGAA 1081
Db 768 ValAlaIleAlaGlyAlaProValThrLeuTrpIlePheThrAspThrGlyTyrThrGlu 787
QY 1082 CGTTATATGGTCACTGACAGAAATGACAGGGCTATTACTAGATCTGTGGCATG 1141

Db 788 ArgTyrMetGlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMet 807
QY 1142 CAACGAGAAAGTCCCTCTCTGAACCAATCGTTTACTGCTCTTACATGGTTTCTCGAT 1201
Db 808 GlnAlaGluLysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAsp 827
QY 1202 GAGAAATCCATTTTCACATACCACATATATTACTGAGTTTATTTAGTCAGGCTGAAAG 1261
Db 828 GluAsnValHisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLys 847
QY 1262 CCATATGATTTACAGATCTATCTCTCAGGAGACACAGCATAGAGTTCTCGAATCGGA 1321
Db 848 ProTyrAspLeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGly 867
QY 1322 GAACTATTGAACGATGATCTTTTGGACTACCTTCAAGAAAACCTTGGATCAGTATGCT 1381
Db 868 GluHisTyrGluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAla 887
QY 1382 GCTCTAAAGTGATA 1396
Db 888 AlaLeuLysValIle 892
RESULT 3
DPP9 HUMAN
ID DPP9 HUMAN STANDARD; PRT; 863 AA.
AC Q86T12; Q6AL137; Q6UAL0; Q6ZMT2; Q6ZNJ5; Q8N2J7; Q8N3F5; Q8WXD8;
AC Q96NT8; Q9BVP3;
DT 13-SEP-2005 (Rel. 48, Created)
DT 13-SEP-2005 (Rel. 48, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Dipeptidyl peptidase 9 (EC 3.4.14.5) (Dipeptidyl peptidase IX) (DPP9)
DE (Dipeptidyl peptidase-like protein 9) (DPLP9) (Dipeptidyl peptidase
IV-related protein 2) (DPRP-2).
GN Name=DPP9; Synonyms=DPRP2;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 1), AND TISSUE SPECIFICITY.
RX MEDLINE=22347328; PubMed=12459266; DOI=10.1016/S0378-1119(02)0059-4;
RA Olsen C., Wagtmann N.;
RT "Identification and characterization of human DPP9, a novel homologue
of dipeptidyl peptidase IV.";
RL Gene 299:185-193(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 1), CATALYTIC ACTIVITY,
RP BIOPHYSICOCHEMICAL PROPERTIES, ENZYME REGULATION, TISSUE SPECIFICITY,
RP AND SUBCELLULAR LOCATION.
RC TISSUE=Colon;
RX PubMed=12662155; DOI=10.1042/BJ20021914;
RA Qi S.Y., Riviere P.J., Trojnar J., Junien J.-L., Akseanva K.O.;
RT "Cloning and characterization of dipeptidyl peptidase 10, a new member
of an emerging subgroup of serine proteases.";
RL Biochem. J. 373:179-189(2003).
RN [3]
RP NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 1), PARTIAL NUCLEOTIDE SEQUENCE
RP [MRNA] (ISOFORM 2), CATALYTIC ACTIVITY, BIOPHYSICOCHEMICAL PROPERTIES,
RP TISSUE SPECIFICITY, AND SUBCELLULAR LOCATION.
RX PubMed=15245913; DOI=10.1016/j.bbexp.2004.03.010;
RA Ajami K., Abbott C.A., McCaughan G.W., Gorrell M.D.;
RT "Dipeptidyl peptidase 9 has two forms, a broad tissue distribution,
cytoplasmic localization and DPV-like peptidase activity.";
RL Biochim. Biophys. Acta 1679:18-28(2004).
RN [4]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 1).
RC TISSUE=Placenta; and Skin;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,
RA Brownstein M.J., Urdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Kzywinaki M.I., Skalek U., Smalhus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [5].
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 30-863 (ISOFORM 3),
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 272-863 (ISOFORM 2), AND
RP PARTIAL NUCLEOTIDE SEQUENCES [LARGE SCALE MRNA] (ISOFORM 5).
RX TISSUE=Glial tumor, Ovary, Spleen, and Trachea;
RC PubMed=14702033; DOI=10.1038/ng1285;
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Makamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA Sekine M., Ohayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,
RA Shiratori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H.,
RA Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E.,
RA Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa K.,
RA Yamazaki M., Ninomiya K., Ishibashi T., Yamashita H., Murakawa K.,
RA Fujimori K., Tanai H., Kimata M., Watanabe M., Hiraoa S., Chiba Y.,
RA Ihida S., Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotchi T.,
RA Kusano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,
RA Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K., Arita M.,
RA Inose N., Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,
RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,
RA Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
RA Hishigaki H., Watanabe T., Sugiyama A., Itakura M., Kawakami B.,
RA Yamazaki M., Watanabe K., Kumagai A., Itakura M., Kawakami B.,
RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ikema Y., Okamoto S.,
RA Kawabata A., Hikiji T., Kobatake N., Inagaki S., Itoh T., Shigeta K., Senba T.,
RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
RA Togaishi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
RA Mizushima-Sugano J., Satoh T., Noguchi S., Komatsu T.,
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
RT "Complete sequencing and characterization of 21,243 full-length human
RT cDNAs";
RL Nat. Genet. 36:40-45 (2004).
RN [6].
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 209-863 (ISOFORM 4), AND
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 298-863 (ISOFORM 2).
RC TISSUE=Melanoma;
RG The German cDNA consortium;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Dipeptidyl peptidases that cleaves off N-terminal
CC dipeptides from proteins having a Pro or Ala residue at position
CC 2.
CC -!- CATALYTIC ACTIVITY: Release of an N-terminal dipeptide, Xaa-Yaa-|-
CC Zaa-, from a polypeptide, preferentially when Yaa is Pro, provided
CC Zaa is neither Pro nor hydroxyproline.
CC -!- ENZYME REGULATION: Inhibited by the serine proteinase inhibitor 4-
CC (2-aminoethyl)benzenesulphonyl fluoride (AEBSF), and by di-
CC isopropyl fluorophosphate.
CC -!- BIOPHYSICOCHEMICAL PROPERTIES:
CC Kinetic parameters:
CC KM=161 uM for Ala-Pro-AMC;
CC KM=180 uM for Ala-Pro-AFC;
CC pH dependence:
CC

Optimum pH is 7.5-8.5. Little activity below pH 6.5;
-!- SUBCELLULAR LOCATION: Cytoplasmic; cytosol.
-!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=5;
CC Name=1; Synonyms=Short; Sequence=Displayed;
CC IsoId=Q86T12-1; Sequence=VSP_013865;
CC Name=2; Synonyms=Long;
CC IsoId=Q86T12-2; Sequence=VSP_013865;
CC Note=Incomplete sequence;
CC Name=3;
CC IsoId=Q86T12-3; Sequence=VSP_013867, VSP_013868;
CC Name=4;
CC IsoId=Q86T12-4; Sequence=VSP_013869;
CC Name=5;
CC IsoId=Q86T12-5; Sequence=VSP_013865, VSP_013866;
CC Note=Incomplete sequence;
CC TISSUE SPECIFICITY: Ubiquitously expressed, with highest levels in
CC liver, heart and muscle, and lowest levels in brain.
CC -!- SIMILARITY: Belongs to the peptidase S9B family. DPPIV subfamily.
CC -!- CAUTION: Ref.6 (CAD39039) sequence differs from that shown due to
CC frameshifts in positions 432 and 460.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC -----
CC EMBL: AF452102; AAL47179.1; -; mRNA.
CC EMBL: AF172660; AAO17262.1; -; mRNA.
CC EMBL: AF542510; AAO73880.2; -; mRNA.
CC EMBL: AV374518; AAO83119.1; -; mRNA.
CC EMBL: BC000570; AAO09070.1; -; mRNA.
CC EMBL: BC037848; AAO37948.1; -; mRNA.
CC EMBL: AK054656; BAC170784.1; ALT_INIT; mRNA.
CC EMBL: AK075050; BAC1362.1; -; mRNA.
CC EMBL: AK131499; BAC8150.1; -; mRNA.
CC EMBL: AL034376; CAD39039.3; ALT_FRAME; mRNA.
CC EMBL: CH627380; CAH10477.1; -; mRNA.
CC MEROPS: S09.019; -;
CC Ensembl: ENSG00000142002; Homo sapiens.
CC HGNC: HGNC:18648; DPP9.
CC MIM: 608258; -;
CC InterPro: IPR001375; Peptidase S9.
CC InterPro: IPR002469; Peptidase S9B.
CC InterPro: IPR000379; Ser. esters.
CC Pfam: PF00930; DPPIV_N; 1.
CC Pfam: PF00326; Peptidase S9; 1.
CC Alternative splicing; Aminopeptidase; Hydrolase; Protease;
CC Serine protease.
CC ACT_SITE 730 730 Charge relay system (By similarity).
CC ACT_SITE 808 808 Charge relay system (By similarity).
CC ACT_SITE 840 840 Charge relay system (By similarity).
CC VARSPPLIC 1 1 M -> LSRVPCVRCRPLPLPGSQSRMSRDEAPLD
CC PGRPAQSRGRTSRVSVCWNGSLDLEGTALLRSAAE
CC RLMEKVKLRDLKENTGSRSPSLNSEGAE (in
CC isoform 2 and isoform 5).
CC /FTid=VSP_013865.
CC /FTid=VSP_013866.
CC /FTid=VSP_013867.
CC /FTid=VSP_013868.
CC /FTid=VSP_013869.
CC /FTid=VSP_013869.
CC I -> N (in Ref. 3).
CC C -> W (in Ref. 5; BAC85150).
CC L -> P (in Ref. 5; BAC18643).
CC G -> C (in Ref. 5; BAC170784).
CC SEQUENCE 863 AA; 98263 MW; 40FE0B78E26CDED5 CRC64;
CC SQ

Alignment Scores:				
Pred. No.:	5.19e-117	Length:	863	
Score:	1494.50	Matches:	281	
Percent Similarity:	62.9%	Conservative:	69	
Best Local Similarity:	50.5%	Mismatches:	105	
Query Match:	50.5%	Indels:	101	
DB:	1	Gaps:	2	
US-10-825-632-6 (1-1669) x DPP9_HUMAN (1-863)				
QY	2	ACAGGTACGACAAATCTTAAGATGTCACCTTTTAAAGATGTCAGAAATATGATGTGCTGAA	61	
DB	308	ThrGlySerLysAsnProLysIleAlaLeuLysLeuAlaGluPheGlnThrAspSerGln	327	
QY	62	GGAGGATCATAGATGTCATAGATAAGGAACCTAAATCAACCTTTTGGAGATCTTCTATTGAA	121	
DB	328	GlyLysIleValSerThrGlnGluLysGluLeuValGlnProPheSerSerLeuPhePro	347	
QY	122	GGAGTGAATATATTCAGAGCTGGATGACTCCTCAGGAAATATGCTTGTGTCATC	181	
DB	348	Ly8ValGluTyrIleAlaArgAlaGlyThrArgAspGlyLysTyrAlaTyrAlaMet	367	
QY	182	CTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAATATTATTC	241	
DB	368	PheLeuAspArgProGlnGlnTrpLeuGlnLeuValLeuLeuProProAlaLeuPheIle	387	
QY	242	CCAGTAGAAGATGATGTATGGAAAGCGAGAGACTCATTGAGTCAGTCCCTGATCTGTG	301	
DB	388	ProSerThrGluAsnGluGlnArgLeuAlaSerAlaArgAlaValProArgAsnVal	407	
QY	302	ACGCCACTAATATCTATGAAGAAACACAGACATCTCGATAAATATCCATGACATCTTT	361	
DB	408	GlnProTyrValValTyrGluGluValThrAsnValTyrIleAsnValHisAspIlePhe	427	
QY	362	CACTTTTCCCAAGTCAC---GAAGAGAAATGTAGTTTATTTTGGCTCTGAATGC	418	
DB	428	TyrProPheProGlnSerGluGlyGluAspGluLeuCysPheLeuArgAlaAsnGluCys	447	
QY	419	AAACAGGTTTCGTCATTTATACAAATTTACATCTATTATTAAGGAAGCAATATAA	478	
DB	448	LysThrGlyPheCysHisLeuTyrLysValThrAlaValLeuLysSerGlnGlyTyrAsp	467	
QY	479	CGATCCAGTGGTGGCTGCTCCAAAGTGATTTCAAGTGTCCTATCAAAAGAGGAGATA	538	
DB	468	TrpSerGluProPheSerProGlyGluAspGluPheLysCysProIleLysGluGluIle	487	
QY	539	GCAATACCAAGTGTGATGGAGATCTTGGCCGCGATGATCTAATATCCAAGTTGAT	598	
DB	488	AlaLeuThrSerGlyGluTrpGluValLeuAlaArgHisGlySerLysIleTrpValAsn	507	
QY	599	GAAGTCAGAGGCTGTATATTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTG	658	
DB	508	GluGluThrLysLeuValTyrPheGlnGlyThrLysAspThrProLeuGluHisIleLeu	527	
QY	659	TACGTAGTCAGTTACGTAAATCTCGAGAGGTGACAAAGCTGACTGACCGTGGCTACTCA	718	
DB	528	TyrValValSerTyrGluAlaAlaGlyGluIleValArgLeuThrProGlyPheSer	547	
QY	719	CATTCTTGTCATCACTGTCATGCTGCTCTTTTATAGTAAGTATAGTACACCAAG	778	
DB	548	HisSerCysSerMetSerGlnAsnPheAspMetPheValSerHisTyrSerSerValSer	567	
QY	779	AAATCCACACTGTGTGCTCCCTTACAGCTATCAAGTCTCAAGTGAAGTCCCACTTGGCAA	838	
DB	568	ThrProCysValHisValTyrLysLeuSerGlyProAspAspProLeuHisLys	587	
QY	839	ACAAGGAATTTTGGCGCACCATTTTGGATTTCAGCAGGTCCTCTTCTCGACTATACCT	898	
DB	588	GlnProArgPheTrpAlaSerMetMetGluAlaAlaSerCysProProAspTyrValPro	607	
QY	899	CCAGAAATTTCTCTTTTGAAGTACTACTGAGTATTTACATCTGAGTATTTTCTACAG	958	

QY	959	CCTCATGATCTACAGCTCGAAGAAATATCTACTGTCTCTTTCATATATGGTGGTCTCT	1018
DB	628	ProHisAlaLeuGlnProGlyLysLysHisProThrValLeuPheValTyrGlyGlyPro	647
QY	1019	CAG-----	1021
DB	648	GlnValGlnLeuValAsnAsnSerPheLysGlyIleLysTyrLeuArgLeuAsnThrLeu	667
QY	1021	-----	1021
DB	668	AlaSerLeuGlyTyrAlaValValIleAspGlyArgGlySerCysGlnArgGlyLeu	687
QY	1021	-----	1021
DB	688	ArgPheGluGlyAlaLeuLysAsnGlnMetGlyGlnValGluIleGluAspGlnValGlu	707
QY	1021	-----	1021
DB	708	GlyLeuGlnPheValAlaGluLysTyrGlyPheIleAspLeuSerArgValAlaIleHis	727
QY	1021	-----	1021
DB	728	GlyTrpSerTyrGlyGlyPheLeuSerLeuMetGlyLeuIleHisLysProGlnValPhe	747
QY	1022	---GTTGCTATTGCTGGGCCCCAGTCACCTCTGTGGATCTTCTATCATACAGGATACAG	1078
DB	748	LysValAlaIleAlaGlyAlaProValThrValTrpMetAlaTyrAspThrGlyTyrThr	767
QY	1079	GAACGTTATATGGTTCACCTGACAGAGTGAACAGGCTATATCTAGGATCTGTGGCC	1138
DB	768	GluArgTyrMetAspValProGluAsnAsnGlnHisGlyTyrGluAlaGlySerValAla	787
QY	1139	ATGCAAGCAGAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGTTCTCTG	1198
DB	788	LeuHisValGluLysLeuProAsnGluProAsnArgLeuLeuIleHisGlyPheLeu	807
QY	1199	GATCAGAAATGTCATTTGCACATACAGTATATATCTAGTGTGTTTGTAGTGGCTCGA	1258
DB	808	AspGluAsnValHisPheHisThrAsnPheLeuValSerGlnLeuIleArgAlaGly	827
QY	1259	AAGCATATGATTTACAGATCTATCTCAGGAGAGACAGACATAGAGTCTCTCAATCG	1318
DB	828	LysProTyrGlnLeuGlnIleTyrProAsnGluArgHisSerIleArgCysProGluSer	847
QY	1319	GGAGAACATATGAATGCTCTTTGCACTCTTCAAGAAACCTT 1366	
DB	848	GlyGluHisTyrGluValThrLeuLeuHisPheLeuGlnGluTyrLeu	863

RESULT 4

DPP9_MOUSE

ID DPP9_MOUSE STANDARD; PRT; 862 AA.

AC Q8BVG4; Q8KAM9; Q8BWT9;

DT 13-SEP-2005 (Rel. 48, Created)

DT 13-SEP-2005 (Rel. 48, Last sequence update)

DT 13-SEP-2005 (Rel. 48, Last annotation update)

DE Dipeptidyl peptidase 9 (EC 3.4.14.5) (Dipeptidyl peptidase IX) (DPP9)

DE (Dipeptidyl peptidase-like protein 9) (DPLP9).

GN Name=Dpp9;

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;

OC Muroidea; Muridae; Murinae; Mus.

OC NCBI_TaxID=10090;

RN [1]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA]: (ISOFORM 1).

RC STRAIN=C57BL/6J; TISSUE=Liver, and Olfactory bulb;

RX	MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
RA	Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA	Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA	Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojibori T.,
RA	Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,


```
QY 2 ACAGGTACAGCAAACTCAAGTCACTTTTAAGATGTCAGAAATAATGATGATCTGAA 61
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
299 ThrGlySerLysAsnProGlnAlaThrIleLysLeuValGluIleLysThrAspGlnGln 318
QY 62 GGAAGG-----
Db |||||
319 GlyArgValSerLeuCysArgLeuLeuLeuCysValPheLeuLeuArgSerAspArgAla 338
QY 68 -----ATCATAGATGTCATAGATAAGAACTAATTCACCT 103
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
339 ValLeuThrCysLeuThrLeuGlnIleValSerThrGlnAspLysGluLeuAlaValPro 358
QY 104 TTTGAGATCTTATTGAAGAGTTGAATATATTGCCAGAGCTGGATGGACTCCTGAGGGA 163
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
359 PheThrSerLeuPheProGlyThrGluThrIleAlaArgValGlyTrpThrSerAspGly 378
QY 164 AAATATGTTGGTCCATCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATC 223
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
379 LysThrGlyTrpAlaAlaLeuLeuLeuAspArgSerGlnArgLysLeuGlnValLeuLeu 398
QY 224 TCACCTGAATATTATCCAGTAGAAGATGTTATGGAAGCAGAGACTCATTTGAG 283
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
399 ProProAlaPhePheValProValThrAspAspProAlaArgArgGlnGluSerLeuGlu 418
QY 284 TCAGTGGCTGATCTGTGACGCCCACTAATTTCTATGAGAAACACACAGATCTGGATA 343
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
419 AlaValProAspArgThrGlnProThrValIleThrGluThrThrAspValTrpIle 438
QY 344 AATATCCATGACATCTTTCATGTTTCCCAAGTCAGAGAGAAATGAGTTTATT 403
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
439 AsnValHisAspIlePheThrProPheValGlnThrAlaGluAspGluPheThrPheIle 458
QY 404 TTTCGCTCTGAAATGCAAAACAGAGTTCCGTCATCTATTATACAAATTTACATCTTTAAAG 463
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
459 TrpValAsnGluSerLysThrGlyPheSerHisLeuThrLysIleThrSerValLeuHis 478
QY 464 -----GAAAGCAATATAAAGCATGACGAGTGGTGGG----- 493
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
479 ProGlyPheHisCysTrpAlaGluAlaThrHisHisThrGluGlyAspProGlnArgArg 498
QY 494 CTGCCTGCT-----CCAAGTGATTTCAAGTGTCTCATCAAAAGAGGAG 535
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
499 IleProAlaValSerThrAspValProProGlyAspPheLysCysAlaValLysGluGlu 518
QY 536 ATAGCAATACCAAGTGTGAATGGGAAGTTCTTGGCCGGCATGGATCTTAATATCAAGTT 595
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
519 IleThrLeuThrSerGlyGluTrpGluValLeuAlaArgHisGlySerLysIleTrpVal 538
QY 596 GATCAAGTCAAGAGCTGGTATATTTGAAGGCACCAAGACTCCCTTTAGAGCATCAC 655
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
539 AsnGluSerSerLysLeuValThrPheGlnGlyThrArgAspThrProLeuGluHisHis 558
QY 656 CTGTACGTAGTACGTAAATCTGGAGAGGTGACAAAGCTGACTGACCGTGGCTAC 715
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
559 LeuThrValValSerThrAspSerProGlyAspValValArgLeuThrLysProGlyPhe 578
QY 716 TCACATCTTGTGTCATCACT----- 736
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
579 SerHisSerCysSerValSerGlnValLysSerLeuGlnSerAspThrPhePheAsn 598
QY 737 -----CAGCACTGTGACTTC 751
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
599 TyrSerSerIleThrLeuProLeuSerLeuSerSerPheIleTrpGlnAsnPheAspPhe 618
QY 752 TTTTATAGTAAATAGTAAACAGAGAAATCCACACTGTGTCTCTTACCACTATCA 811
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
619 PheValSerHisThrSerValCysThrProProCysValHisValThrLysLeuAsn 638
QY 812 AGTCCTGAAGATGACCACTTCGCAAAACAAAGGAATTTTGGGCCACCATTTGGATCA 871
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
639 SerSerGluSerAspProLeuHisIleValProGluPheTrpAlaSerMetMetGluSer
```

```
QY 872 GCAGGTCTCTTCTCTGACTATACCTCCAGAAATTTCTCTTTCAAAAGTACTACTGA 931
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
659 SerGlyCysProGlyAspThrSerProGluIlePheAspPheGlnGlyLysSerGly 678
QY 932 TTTACATTTGATGGATGCTCTACAGCTCATGATCTACAGCTCGAAAGAAATATCTCT 991
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
679 PheGlnLeuThrGlyMetValThrLysProHisSerLeuGlnProGlyArgLysHisPro 698
QY 992 ACTGTGCTGTTTCATATATATGTGTGCTCTCAGGTT----- 1024
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
699 ThrValLeuPheValThrGlyGlyProGlnValGlnLeuValAsnAsnSerPheLysGly 718
QY 1024 ----- 1024
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
719 MetLysThrLeuArgLeuAsnThrLeuAlaSerLeuGlyTyrrAlaValValIleAsp 738
QY 1024 ----- 1024
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
739 GlyArgGlySerCysGlnArgGlyLeuGluPheGluSerAlaLeuLysAsnLysMetGly 758
QY 1024 ----- 1024
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
759 GlnValGluIleGluAspGlnValGluGlyLeuGlnTyrrValAlaGluLysPheAsnPhe 778
QY 1024 ----- 1024
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
779 ValAspLeuSerArgValAlaIleHisGlyTrpSerTyrrGlyGlyPheLeuSerLeuMet 798
QY 1025 -----GCTATTCTGGGGCCCGCTCACTCTG 1051
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
799 GlyLeuIleGlnArgProAsnValPheLysLeuAlaIleAlaGlyAlaProValThrVal 818
QY 1052 TGGATCTTCTATGATACAGGATACAGGACGTTATATGGTCACCTGACCAGAAATGAA 1111
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
819 TrpMetAlaThrAspThrGlyThrGluArgTyrrMetAspValProGluAsnAsnGln 838
QY 1112 CAGGCTATTACTTATGATCTGTGGCATCGAAGCAGAAAGTTCCTCTCGAACCAAT 1171
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
839 GlnGlyThrGluGluGlySerValAlaLeuHisValAspLysLeuProSerGluProAsn 858
QY 1172 CGTTTACTGCTTTPACATGCTTCTGGATGAGAATGTCATTTTGCACATACCAATATA 1231
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
859 ArgLeuLeuIleLeuHisGlyPheLeuAspGluAsnValHisPhePheHisThrAsnPhe 878
QY 1232 TTACTGAGTTTTTTAGTGAGGCTGGAAAGCCATATGATTTACAGATCTATCTCCAGGAG 1291
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
879 LeuValSerGlnIleIleArgAlaGlyLysProTyrrGlnLeuGlnValThrProAsnGlu 898
QY 1292 AGACACAGCATAGAGTTCTGTAATCGGAGAACATTAATGAACCTGATCTTTGCACATAC 1351
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
899 ArgHisSerIleArgCysProGluSerGlyGluHisTyrrGluIleMetLeuLeuHisPhe 918
QY 1352 CTTCAAGAAACACCTT 1366
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
919 LeuGlnGlnTyrrLeu 923
QY 923 -----
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
RESULT 6
Q6GR22_XENLA PRELIMINARY; PRT; 847 AA.
AC Q6GR22_
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE MGC81312 (TrEMBLrel. 27, Last annotation update)
DE MGC81312 Protein.
GN Name=MGC81312;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus; Xenopus.
OX NCBI_taxid=8355;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Embryo;
```



```
Db 732 LysValAlaIleAlaGlyAlaProValThrLeuTrpMetAlaTyrAspThrGlyTyrThr 751
QY 1079 GAACGTTATATGGTCAACCTGACAGAAATGACAGGGCTATTACTTAGGATCTGGCC 1138
Db 752 GluArgTyrMetGluThrProGluThrAsnGlnMetGlyTyrGluAlaGlySerAlaAla 771
QY 1139 ATGCAAGCAGAAAGTCCCTCGAACCAATCGTTACTCTCTTACATGTTTCTGTG 1198
Db 772 LeuGlnValAspLysLeuProAsnGluProTyrArgLeuLeuIleLeuHisGlyPheLeu 791
QY 1199 GATCAGATGTCATTTTGCACATACACAGTATATTAATCTAGTGTCTTTAGTGAGGCTGGA 1258
Db 792 AspGlnValHisPhePheHisThrAsnPheLeuLeuSerGlnLeuIleArgAlaGly 811
QY 1259 AAGCATATGATTTACAGATCTATCTCAGGAGAGACACAGACATTAAGAGTTCTCGAATCG 1318
Db 812 LysProTyrGlnLeuGlnIleTyrProAsnGluArgHisSerIleArgCysProGluSer 831
QY 1319 GGAAACATATTAAGACTGTCATTTTGGCACTACCTTCAAGAAACCTT 1366
Db 832 GlyGluHisTyrGluIleThrLeuLeuHisPheLeuGlnGluHisLeu 847

RESULT 7
O75273 HUMAN
ID O75273 HUMAN PRELIMINARY; PRT; 508 AA.
AC O75273;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE R26984.1 (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
[1]
NP NUCLEOTIDE SEQUENCE.
RA Lamerdin J.E., McCready P.M., Skowronski E., Adamson A.W.,
RA Burkhardt-Schulz K., Gordon L., Kyle A., Ramirez M., Stillwagen S.,
RA Phan H., Velasco N., Do L., Regala W., Terry A., Barnes J.,
RA Danganan L., Poundstone P., Christensen M., Georgescu A., Avila J.,
RA Liu S., Attix C., Andreise T., Frankheim M., Anico-Keller G.,
RA Coefield J., Duarte S., Lucas S., Bruce R., Thomas P., Quan G.,
RA Kronmiller B., Arellano A.S., Montgomey M., Ow D., Nolan M., Trong S.,
RA Kobayashi A., Olsen A.S., Carrano A.V.;
RA Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC005594; AAC33801.1; -; Genomic_DNA.
DR MEROPS; S09.019; -.
DR Ensembl; ENSG00000142002; Homo sapiens.
DR GO; GO:0003824; F.catalytic activity; IEA.
DR InterPro; IPR001375; Peptidase_S9.
DR InterPro; IPR002469; Peptidase_S9B.
DR InterPro; IPR000379; Ser_estr.
DR Pfam; PF00930; DPPIV_N; 1.
DR Pfam; PF00326; Peptidase_S9; 1.
FT NON_TER 1
FT 1
SQ SEQUENCE 508 AA; 57750 MW; 2F4D6645BE2D2C89 CRC64;

Alignment Scores
Pred. No.: 3,75e-91 Length: 508
Score: 1187.50 Matches: 241
Percent Similarity: 56.2% Conservative: 64
Best Local Similarity: 44.4% Mismatches: 93
Query Match: 40.2% Indels: 145
DB: 2 Gaps: 7

US-10-825-632-6 (1-1669) x O75273_HUMAN (1-508)
QY 68 ATCATAGATGTCATAGATAAGCACTAATTCACCTTTTGGATTTCTATTGAGGAGTT 127
Db 1 IleValSerThrGlnGluSerGlnLeuValGlnProPheSerLeuPheProLysVal 20
QY 128 GAATATATGGCCAGAGCTGGATGGACTCTCTGAGGGAAAAATATGCTTGGTCCATCTACTA 187
```

```
Db 21 GluTyrIleAlaArgAlaGly-----AlaTrpAlaMetPheLeu 33
QY 188 GATCGCTCCAGACTCCCTTACAGATAGTGTGATCTCACCTGAATATTATTATCCAGTA 247
Db 34 AspArgProGlnGlnTrpLeuGlnLeuValLeuLeuProAlaLeuPheIleProSer 53
QY 248 GAAGATGATGTATGTAAGGAGAGAGACTCAATTGAGTCAGTCGCTGATCTCTGTGAGCCA 307
Db 54 ThrGluAsnGluGluGlnArgLeuAlaSerAlaArgAlaValProArgAsnValGlnPro 73
QY 308 CTAATTATCTATGAAGAAACACACAGACATCTGATATAATATCATGACATCTTTTCATGTT 367
Db 74 TyrValValTyrGluGluValThrAsnValTrpIleAsnValHisAspIlePheTyrPro 93
QY 368 TTTCCCAAAAGTACAC---GAAGAGGAAATTTGAGTTTATTATTTTGCCTCTGAATGCAAAACA 424
Db 94 PheProGlnSerGluGlyGluAspGluLeuCysPheLeuArgAlaAsnGluCysLysThr 113
QY 425 GGTTCCTCGTCATTTATACAAAATTTACATCTATTTTAAAGGAAAGCAAAATATAAACGATCC 484
Db 114 GlyPheCysHisLeuTyrLysValThrAlaValLeuLysSerGlnGlyTyrAspTrpSer 133
QY 485 AGTGGTGGCTGCTGCTCCCAAGTGATTTCAAGTGTCTTATCAAGAGAGAGATACGAATT 544
Db 134 GluProPheSerProGly-----GluGlyGluGlnSerLeu 145
QY 545 ACCAGTGGTGAATGGGAGTCTTGGCCGCGATGATCTAATATCAAGTTGATGAAGTC 604
Db 146 ThrAsnAlaIleTrp-----ValAsnGluGlu 154
QY 605 AGAAGGTGATATATTGTAAGGCACCAAGACTCCCTTTAGACATCACTGTACAGTA 664
Db 155 ThrLysLeuValTyrPheGlnGlyThrLysAspThrProLeuGluHisLeuTyrVal 174
QY 665 GTCAAGTTACGTAATCTCGAGAGGTGACACAGCTGACTGACCGGTGCTACTCATCTCT 724
Db 175 ValSerTyrGluAlaAlaGlyGluIleValArgLeuThrThrProGlyPheSerHisSer 194
QY 725 TGCTGCATCAGTCAGCACTGTGACTCTTTTATAAGTAAGTATAGTAACACAGAAATCCA 784
Db 195 CysSerMetSerGlnAsnPheAspMetPheValSerHisTyrSerSerValSerThrPro 214
QY 785 CACTGTGTGTCCTTTTACCAAGCTATCAAGTCTCTGAAGATGACCCCAACTTGCAAAACAAAG 844
Db 215 ProCysValHisValTyrLysLeuSerGlyProAspAspProLeuHisLysGlnPro 234
QY 845 GAATTTTGGGCCACCATTTTGGATTGAGTTCAGTGTATGGGATGCTCTCAAGCCTCAT 904
Db 235 ArgPheTrpAlaSerMetMetGluAlaAla-----Lys 245
QY 905 ATTTTCTCTTTTGAAGTACTACTGGATTTACATTTGATGGGATGCTCTCAAGCCTCAT 964
Db 246 IlePheHisPheHisThrArgSerAspValArgLeuTyrGlyMetIleTyrLysProHis 265
QY 965 GATCTACAGCTGGAAGAAATATCTACTGTGCTGTTTCATATATGGTGGTCTCAG--- 1021
Db 266 AlaLeuGlnProGlyLysLysHisProThrValLeuPheValTyrGlyGlyProGlnVal 285
QY 1021 ----- 1021
Db 286 GlnLeuValAsnAsnSerPheLysGlyIleLysTyrLeuArgLeuAsnThrLeuAlaSer 305
QY 1021 ----- 1021
Db 306 LeuGlyTyrAlaValValIleAspGlyArgGlySerCysGlnArgGlyLeuArgPhe 325
QY 1021 ----- 1021
Db 326 GluGlyAlaLeuLysAsnGlnMetGlyGlnValGluIleGluAspGlnValGluGlyLeu 345
QY 1021 ----- 1021
```

```
Db 346 GlnPheValAlaGluLysTyrGlyPheIleAspLeuSerArgValAlaIleHisGlyTrp 365
QY 1022 -----GTT 1024
Db 366 SerTyrGlyGlyPheLeuSerLeuMetGlyLeuIleHisLysProGlnValPheLysVal 385
QY 1025 GCTATTGTCGGGGCCCGAGTCTACTCTGTGGATCTTCTATACATACAGGATACAGGACGT 1084
Db 386 AlaIleAlaGlyAlaProValThrValTrpMetAlaTyrAspThrGlyTyrThrGluArg 405
QY 1085 TATATGGTTCACCTCCACCAAGTAAACAGAGCTATTACTTAGGATCTGTGGCCATGCCAA 1144
Db 406 TyrMetAspValProGluAsnAsnGlnHisGlyTyrGluAlaGlySerValAlaLeuHis 425
QY 1145 GCAGAAAGTTCCTCTGCAACCAATCGTTTACTGCTTTCATACATGCTTTCCTGGATGAG 1204
Db 426 ValGluLysLeuProAsnGluProAsnArgLeuLeuIleHisGlyPheLeuAspGlu 445
QY 1205 AATGTCATTTTGCATACACAGTATATTAATCTAGATTTTATGAGGGTGGGAAAGCCA 1264
Db 446 AsnValHisPhePheHisThrAsnPheLeuValSerGlnLeuIleArgAlaGlyLysPro 465
QY 1265 TATGATTTA-----CAGATCTATCTCAGGAGACAC 1297
Db 466 TyrGlnLeuGlnValAlaLeuProValSerProGlnIleTyrProAsnGluArgHis 485
QY 1298 AGCATAGAGTTCCTGATCGGAGACATTAATGACTGCTTTTGCACCTACCTTCAA 1357
Db 486 SerIleArgCysProGluSerGlyGluHisTyrGluValThrLeuLeuHisPheLeuGln 505
QY 1358 GAAACCTT 1366
Db 506 GlutyrLeu 508

RESULT 8
ID Q7PTT8_ANOGA PRELIMINARY; PRT; 621 AA.
AC Q7PTT8;
DT 01-MAR-2004 (T:EMBLrel. 26, Created)
DT 01-MAR-2004 (T:EMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (T:EMBLrel. 26, Last annotation update)
DE ENSANGP0000015447 (Fragment).
GN ORFNames=ENSANG0000012958;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Ehdopterygota; Diptera; Nematocera; Culicoidea; Culicidae;
OC Anophelinae; Anopheles.
OX NCBI_TaxID=180454;
[1]
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=PEST;
RC The Anopheles gambiae Sequence Committee;
RG "Anopheles gambiae re-annotation.";
RT Submitted (Apr-2002) to the EMBL/GenBank/DDBJ databases.
[2]
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=PEST;
RC The Anopheles gambiae Sequence Committee;
RG Submitted (Apr-2004) to the EMBL/GenBank/DDBJ databases.
RL Submitted (Apr-2004) to the EMBL/GenBank/DDBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AA001008287; EAA03335.3; -; Genomic_DNA.
DR MEROPS; S09.016; -.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001375; Peptidase_S9.
DR InterPro; IPR002469; Peptidase_S9B.
DR InterPro; IPR000379; Ser_estrs.
DR Pfam; PF00930; DPPIV N; I.
DR Pfam; PF00326; Peptidase_S9; 1.
NT NON_TER 1
```

```
FT NON_TER 621
SQ SEQUENCE 621 AA; 70546 MW; F6EA8463A343BBF3 CRC64;
Alignment Scores:
Pred. No.: 2,42e-66 Length: 621
Score: 894.00 Matches: 211
Percent Similarity: 46.1% Conservative: 69
Best Local Similarity: 34.8% Mismatches: 147
Query Match: 30.2% Indels: 180
DB: Gaps: 14
US-10-825-632-6 (1-1669) x Q7PTT8_ANOGA (1-621)
QY 5 GTTACAGCAAAATCCTTAAGTCACTTTTAAGATGTCAGAAATATGATTGATGCTGAAGGA 64
Db 41 GlyThrProAsnAlaLysSerLysLeuLysLeuValGlnPheArgLeuSerGluAsnLeu 60
QY 65 AGGATCATAGATCATAGATAAGAACTAATTAACCTTTTGAGATTCTTATTGAAAGGA 124
Db 61 ArgIleThrAspValCysIleLysGluLeuGlnCysProLeuThrPheAlaPheProTrp 80
QY 125 GTTGAATATATTTGCCAGAGCTGGATGACTCTCTAGAGGAAATATGCTGTGTCATCCTA 184
Db 81 LeuGluTyrIleValArgValGlyTrpProAspSerArgTyrValTrpAlaGlnLeu 100
QY 185 CTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAATATTATTATCCCA 244
Db 101 LeuAspArgProGlnArgLeuGluLeuValLeu-----LeuPro 114
QY 245 GTAGAGATCATGTTTATGAAAGGAGAGACTCATTGAGTCAGTGCCT----- 292
Db 115 Val---AspAsnPheCysGluIleTyrSerSerSerSerLeuProProArgLysSer 133
QY 293 -----GATTCTGTGACGCCACTA---ATTATCTAT 319
Db 134 SerSerGlyTrpArgSerProLeuAspLysSerThrThrArgProLeuGlnValIleTyr 153
QY 320 GAAGAAACAACAGACATCTGGATAATATCCATGACATCTTTTCATGTTTTTCCCAAAAGT 379
Db 154 ThrGluThrSerSerSerTrpValAsnValHisAspValLeuGln---PheValGluLeu 172
QY 380 CACGAAGAGAAATGAGTTTATTTTTCCTCTGATGCAAAACAGGTTTCGTCATTTA 439
Db 173 SerGluGlnGluValThrPheLeuTrpAlaSerGlu---GluSerGlyPheArgHisLeu 191
QY 440 TACAAAATTACATCTATTTTAAAGGAAACAATATAACGATCCAGTGGTGGGCTGCCT 499
Db 192 TyrLeuValThrSerSerLeuSerProAsnGluVal-----SerSerGlyGlyAlaGly 209
QY 500 GCTCCAAGTGTATTTCAAGTGTCTCT----- 523
Db 210 Ala---ThrAspHisSerLeuProSerMetAlaCysIleGlySerThrLeuValAlaArg 228
QY 524 ATCAAGAGGAGATAGCAATTACCAGTGTGAATGGGAAGTTCTTTGGCGGCGATGATCT 583
Db 229 IleValGlnLysValThrLeuThrGlyGlyAspTrpGluValLeuGlyArg----- 245
QY 584 AATATCCAAGTTGATGAAGTCAGAGCTGTATATTTTGAAGGACCAACAGACTCCCT 643
Db 246 AsnValTrpTyrAspArgValArgGlnLeuValTyrPheMetGlyLeuArgGluThrPro 265
QY 644 TTAGAGATCACCTGTAGTACGTACGTACGTAAATCTCTGGAGAGGTGACAGGCTGACT 703
Db 266 LeuGluLysHisLeuTyrValValSerLeuAlaGlnProAsnGlnLeuArgLeuLeuThr 285
QY 704 GACCTGGCTACTCACATTTCTGCTGATCAGTCAGCTGCTGACTCTTTTATAAGTAAG 763
Db 286 MetProGlyTyrSerPheThrValGluPheAsnAspAspCysThrLeuPheLeuGlnThr 305
QY 764 TATAGTAACACAGAAAGATCCACACTGTGTGTCCTCTTTTACAAGCTA-----TCAAGT 814
Db 306 TyrCysAsnIleSerThrLeuProSerTrpGluLeuValArgIleAlaHisAspSerAsn 325
```


Db	188	GlyThrProAsnAlaLysSerLysLeuValGlnPheArgLeuSerGluAsnLeu	207
Qy	65	AGGATCATAGTCATAGTAAGGAACCTAACTTTTGAGATTCTATTGGAAGGA	124
Db	208	ArgGlyThrAspValCysLeuLysGluGlnCysProLeuThrPheAlaPheProTyr	227
Qy	125	GTTGAATATATGCCAGAGCTGGATGGACTCTCTGAGGGAATAATATGCTTGGTCCATCCTA	184
Db	228	LeuGluTyrTrileValArgValGlyTyrThrProAspSerArgTyrValTrpAlaGlnLeu	247
Qy	185	CTAGATCGCTCCAGACTCGCTTACAGATAGTGTGATCTCACCTGAAATATTATATCCCA	244
Db	248	LeuAspArgProGlnArgLeuGluLeuValLeuLeuProValAspAsnPheCysGlu	267
Qy	245	GTAGAAGATGATGTTATGAAAGGCAGAGACTCATTGAGTCAGTGCCT	292
Db	268	IleTyrSerSer	299
Qy	293	-----GATTCTGTGACGCCACTA---ATTATCTATGAAGAAACAACAGACATCTGG	340
Db	280	HisArgLysSerThrThrArgProLeuGlnValIleTyrThrGluThrSerSerSerTyr	299
Qy	341	ATAAATATCCATCACATCTTTCATGTTTTCCTCCCAAGTCACGAGAGGAATATGAGTTT	400
Db	300	ValAsnValHisAspValLeuGln---PheValGluLeuSerGluGlnValThrPhe	318
Qy	401	ATTTTGGCTCTCAATGCAAAACAGGTTTCCTGCTATTATACAAATATACATCTATTTTA	460
Db	319	LeuTrpAlaSerGlu---GluSerGlyPheArgHisLeuTyrLeuValThrSer	335
Qy	461	AAGGAAGCAAAATATAACAGATCCAGTGTGGCTGCTCTCCAGTGATTTCAAGTGT	520
Db	336	-----LysValSerSerGlyValGlyAla-----ThrAspHisSerLeu	349
Qy	521	CCT-----ATCAAAGAGGAGATAGCAATT	544
Db	350	ProSerMetAlaCysIleGlySerThrLeuValAlaArgIleValGlnLysValThrLeu	369
Qy	545	ACCAGTGGTGAATGGGAAGTCTTGGCGGCATGGATCTAATATCCAAAGTTGATGAAGTC	604
Db	370	ThrGlyGlyAspTrpGluValLeuGlyArg-----AsnValTrpTyrAspArgAla	386
Qy	605	AGAAGGTGTATATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACCTGTACGTA	664
Db	387	ArgGlnLeuValTyrPheMetGlyLeuArgGluThrProLeuGluLysHisLeuTyrVal	406
Qy	665	GTCAAGTTACGTAATCTCGAGAGGTGACAAAGCTGACTGACCGCTGGCTACTCACATTCT	724
Db	407	ValSerLeuAlaGlnProAsnGlnLeuArgLeuLeuThrMetProGlyTyrSer	424
Qy	725	TGCTGCATCAGTCAGCACTGTGACTTCTTTATAAGTATAGTAACT	772
Db	425	-----PheThrValGluPheAsnAspValSerValPhe	435
Qy	773	CAGAAGAATCCACACTGTGTCCCTTTTACAGCTATCAAGTCTCGAAGATGACCCCACT	832
Db	436	GluArgLysGlyAlaCysProAsnThrAlaAsnGlyAsnGlyCysSerHisGlyProThr	455
Qy	833	TGCNAACAAGAAATTTGGGCCACCATTTGGATTTCAGCAGCTCTCTTCTCGACTAT	892
Db	456	ProProProThrProIleAspAlaLeuArgLeuCysSerValGlyTyrLeuThrGluGly	475
Qy	893	ACTCCT-----CCAGAAATTTCTCTTTTGAAGTACTACTCGGA	931
Db	476	GlyProSerGluAsnThrGlnTyrAsnProSerIleHisSerProGlnIleSerSerGly	495
Qy	932	TTTACATGTATGGATGCTTACAGGCTCATGATCTACAGCTGGGAAGAAATATCCT	991
Db	496	AspValLeuTyrAlaMetValPheLysProHisAsnPheMetLeuGlyValLysTyrPro	515
Qy	992	ACTGTGCTGTATATATCGTGTCTCAG-----	1021
Db	516	ThrValLeuAsnValTyrGlyGlyProGluValGlnThrValSerAsnThrPheLysGly	535
Qy	1021	-----	1021
Db	536	MetArgGlnLeuArgMetHisMetLeuAlaSerGlnGlyTyrCysValIleCysValAsp	555
Qy	1021	-----	1021
Db	556	SerArgGlySerArgHisArgGlyValGluPheGluSerTyrIleArgCysArgMetGly	575
Qy	1021	-----	1021
Db	576	ThrValGluLeuSerAspGlnValGluValLeuArgIleLeuAlaAspGlnLeuGlyTyr	595
Qy	1021	-----	1021
Db	596	IleAspMetAspArgValAlaIleHisGlyTyrSerTyrGlyTyrLeuSerLeuMet	615
Qy	1022	-----GTTGCTATTGCTGGGGCCCCAGTCACTCTG	1051
Db	616	GlyLeuValGlnTyrProGluIlePheLysValSerIleAlaGlyAlaProValThrSer	635
Qy	1052	TGGATCTTCTATCATACAGGATACACGGAACCTTATATATGGTCCACCTGCAGCAATGAA	1111
Db	636	TrpGluTyrTyrAspThrGlyTyrThrGluArgTyrMetAspLeuProAspSerAsnArg	655
Qy	1112	CAGGCTATTACTTACGATCTGTGGCCATGCAAGCAGAAAAGTCCCTCTGAACCAAT	1171
Db	656	SerGlyTyrAlaAlaGlySerValLeuAsnTyrIleGlnLysPheProAspGluAspAsn	675
Qy	1172	CGTTACTGCTCTTACATGGTTTCTGGATGAGATGTCATTTTGCATACCATGATATA	1231
Db	676	ArgLeuLeuIleIleHisGlyLeuIleAspGluAsnValHisPheHisThrSerGln	695
Qy	1232	TTACTGATGTTTTTAGTAGGGCTGGAAAGCCATATGATTTACAGATCTATCTCCAGAG	1291
Db	696	LeuValSerArgLeuValArgAlaAsnLysProTyrGlnLeuGlnValTyrProAsnGlu	715
Qy	1292	AGACACAGCATAAGATGTTCTGAAATCGGGAACATATATGAACCTCATCTTTGCACCTAC	1351
Db	716	ArgHisSerLeuArgAsnLeuGluAlaSerLysHisTyrGluThrLysLeuSerPhe	735
Qy	1352	CTTCAAGAAACCTT	1366
Db	736	LeuGlnAsnHisLeu	740
RESULT 11			
Q9VC20 DROME			
ID	Q9VC20	DROME PRELIMINARY;	PRT: 1053 AA.
AC	Q9VC20	Q8IH07;	
DT	01-MAY-2000	(TrEMBLrel. 13, Created)	
DE	10-MAY-2000	(TrEMBLrel. 13, Last sequence update)	
DE	10-MAY-2000	(TrEMBLrel. 30, Last annotation update)	
GN	CG3744-PA	isoform A (CG3744-pc, isoform c) (LD33755p).	
OR	ORFNames=CG3744;		
OS	Drosophila melanogaster	(Fruit fly).	
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;		
OC	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;		
OC	Ephydroidea; Drosophilidae; Drosophila.		
OX	NCBI_taxid=7227;		
RN	[1]		
NUCLEOTIDE SEQUENCE.			
RX	MEDLINE=20196006;	PubMed=10731132;	DOI=10.1126/science.287.5461.2185;
RA	Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,		
RA	Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,		
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,		
RA	Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,		
RA	Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,		
RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,		
RA	Abriel J.F., Agbayani A., An H.-J., Andrews-Bianknoch C., Baldwin D.,		
RA	Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,		
RA	Besson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,		
RA	Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,		
RA	Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,		

Db 631 ValThrAlaSerLeuLeuLeuSerGln-----AlaAsnGlyGlnProAspPro 646
QY 506 -----AGTGATTTCAAGTGCCTATCAAA 529
Db 647 GlySerValGlyAlaGlnProSerPheValAspLeuSerAlaLeuGlnProArgIleLeu 666
QY 530 GAGGAGATACGAATTACCTAGTGTGTAATGGAGTTCTTGGCGGCATGGATCTAATATC 589
Db 667 AsnLysValAlaLeuThrSerGlyGluTrpGluValLeuAlaArg-----AsnLeu 683
QY 590 CAGTTTCATCAAGTCAGAGCGTGTATATTTGAAGCACCAAGACTCCCTTTTAGAG 649
Db 684 TrpValAspLysAlaAsnLysLeuValTyrPheValGlyLeuArgAspThrProLeuGlu 703
QY 650 CATCACTGTAGTGTACGTACGTAAATCTCGAGAGGTGCACAAAGCGCTGACTGACCGT 709
Db 704 LysHisLeuTyrValValSerLeuGluArgProGluHisIleArgLeuLeuThrGluPro 723
QY 710 GGTCTACTCATTCTTTGC-----TGCACTAGT 736
Db 724 GlyTyrSerTyrLeuValGluPheAspAspHisPheAsnAspLeuPheProIlePheSer 743
QY 737 CAGCACTGTGACTTCTTTTATAGTATAGTAAC---CAGAAAGATCCACACTGTGTG 793
Db 744 GlnGlnCysLysLeuMetLeuLeuValTyrCysAsnIleGlnArgLeuProSerCysLys 763
QY 794 TCCCTTTACAAGCTATCAAGTCTCGAAGATGACCCCACTTCGAAACACAAAGGAATTTGG 853
Db 764 ValMet-----ArgValAsnGlnThrCysSerAsnGlyGlyValAsn 777
QY 854 GCCACCAATTTGGATTCAGCAGGT-----CCTCTCTCTGACTAT 892
Db 778 GlyIleGlnIleSerLeuValGlyTyrLeuHisGluGlyGlyLysProGluProGlnTyr 797
QY 893 ACTCTCCAGAAATTTCTTTTGAAGTACTACTGATTTACATTTGATGGATGCTC 952
Db 798 Cys---ProGlnIlePheSerProGlnLeuProSerGlyAspIleValTyrAlaMetVal 816
QY 953 TACAGCGCTCATGCTACAGCTCGAAGAAATATCTACTGTGCTGTCATATAGT 1012
Db 817 PheLysProHisAsnPheGluLeuGlyValLysTyrProThrValLeuAsnValTyrGly 836
QY 1013 GGTCTCTCAG----- 1021
Db 837 GlyProGluValGlnThrValAsnAsnThrPheLysGlyLysHisGlnLeuArgMetHis 856
QY 1021 ----- 1021
Db 857 MetLeuAlaAlaGlnGlyTyrCysValIleCysIleAspSerArgGlySerArgHisArg 876
QY 1021 ----- 1021
Db 877 GlyLysArgPheGluSerHisIleArgGlyArgMetGlyGlnValGluLeuThrAspGln 896
QY 1021 ----- 1021
Db 897 ValAspAlaLeuArgSerLeuSerAspGlnLeuGlyTyrIleAspMetAspArgValAla 916
QY 1021 ----- 1021
Db 917 IleHisGlyTrpSerTyrGlyTyrLeuSerLeuMetGlyLeuValGlnTyrProLys 936
QY 1022 -----GTTGCTATTCTGGGCGCCAGTCACTCTGGATCTCTATGATACAGA 1072
Db 937 IlePheLysValAlaIleAlaGlyAlaProValThrAsnTyrGluTyrTyrAspThrGly 956
QY 1073 TACACGAGCGTTATATGGTGCACCTGACAGATGAACAGGCGTATTACTTAGATCT 1132
Db 957 TyrThrGluArgTyrMetAspMetProGlnAsnAsnGluAlaGlyTyrSerAlaGlySer 976
QY 1133 GTGGCATGCAAGCAGAAAGTTCCCTCTCGAACCAATCGTTTACTGCTCTTACATGTT 1192
Db 977 ValLeuGluTyrValAsnSerPheProGluGluAspLysArgLeuLeuLeuIleHisGly 996

QY 1193 TTCCTGGATGATGCTCATTTTGCATCTTTCGACATACAGTATATTACTAGTITTTTACTGAGG 1252
Db 997 LeuIleAspGluAsnValHisPheCysHisThrSerArgLeuIleSerAlaLeuAsnLys 1016
QY 1253 GCTGGAAGCCATATGATTTTACAGATCTCTCTCAGGAGAGACACAGCATAGAGATTCCT 1312
Db 1017 AlaAsnLysProTyrGluValHisLeuPheProGluGluArgHisSerLeuArgAsnLeu 1036
QY 1313 GAATCGGAGAGAACATATTGAATCGCATCTTTTCGACTACTCTTCAA 1357
Db 1037 GluSerAsnLysAsnTyrGluThrLysLeuLeuSerPheLeuGln 1051
RESULT 12
Q9VC19_DROME PRELIMINARY; PRT; 1113 AA.
AC Q9VC19;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE CG3744-PB, isoform B.
GN Name=CG3744; ORFNames=CG3744;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gockayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Wan K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Miklos G.L.G.,
RA Abell J.F., Agbayani A., An H.-J., Andrews-Fianknoch C., Baldwin D.,
RA Bailett R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bertram B.P., Brokstein P., Brottier P.,
RA Borokova D., Botchan M.R., Bouck J., Braketein D., Bolshakov S.,
RA Chertys K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Curry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA De Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fogle C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laslo P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson J.,
RA Merkulov G., Milshina N.V., Mobarry C., Murphy L., Muzny D.M., Nelson D.L.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Paclet J.M.,
RA Patzold M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpsons R.D.C., Scheeler F., Shen H.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195 (2000).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426065; PubMed=12537568;
RA Celinker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Friese E., Hodgson A.,

Db 917 MetLeuAlaAlaGlnGlyTyrCysValIleCysIleAaspSerArgGlySerArgHisArg 936
Qy 1021 -----
Db 937 GlyTyrArgPheGluSerHisIleAlaGlyArgMetGlyGlnValGluLeuThrAspGln 956
Qy 1021 -----
Db 957 ValAspAlaLeuArgSerLeuSerAspGlnLeuGlyTyrIleAaspMetAspArgValAla 976
Qy 1021 -----
Db 977 IleHisGlyTrpSerTyrGlyTyrLeuSerLeuMetGlyLeuValGlnTyrProLys 996
Qy 1022 -----GTTCTATTCTGGGCCCGCCAGTCATCTGTGGATCTTCTATGATACAGGA 1072
Db 997 IlePheLysValAlaIleAlaGlyAlaProValThrAspThrPgluTyrTyrAspThrGly 1016
Qy 1073 TACACGAGAGCTTATATGGTCACTGACCGTACAGATGAACAGGGCTATTACTTAGGATCT 1132
Db 1017 TyrThrGluArgTyrMetAspMetProGlnAenAenGluAlaGlyTyrSerAlaGlySer 1036
Qy 1133 GTGCCCATGCAACGAGAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGGT 1192
Db 1037 ValLeuGluTyrValAenSerPheProGluGluAaspLysArgLeuLeuIleHisGly 1056
Qy 1193 TTCTGTGATGAGATGCTCATTTTGCACATACCATGATATATTACTAGTTTCTTTAGTGAGG 1252
Db 1057 LeuIleAaspGluAenValHisPheCysHisThrSerArgLeuIleSerAlaLeuAenLys 1076
Qy 1253 GCTGGAAAGCCATATGATTACAGATATCTCTCAGGAGACACACAGATAGAGTTTCTTCA 1357
Db 1077 AlaAenLysProTyrGluValHisLeuPheProGluGluArgHisSerLeuArgAenLeu 1096
Qy 1313 GAATCGGAGAACATPATGAATGCTCTTTTGCATCTCTTCA 1357
Db 1097 GluSerAenLysAenTyrGluThrLysLeuSerPheLeuGln 1111
RESULT 13
QSTXJ2_ANOGA
ID QSTXJ2 ANOGA PRELIMINARY; PRT; 557 AA.
AC QSTXJ2;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE ENSANGP00000029249 (Fragment).
GN ORFNames=ENSANG0000012958;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Culicidae;
OC Anophelinae; Anopheles.
OX NCBI_TaxID=180454;
RN (1)
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
RT "Anopheles gambiae re-annotation";
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
RN (2)
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
CC -! CAUTION, The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAAB01008287; EAL42106.1; -; Genomic DNA.
DR GO; GO:0003824; P:serine-type peptidase activity; IEA.
DR GO; GO:0008233; F:serine-type peptidase activity; IEA.
DR GO; GO:0006535; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR0011375; Peptidase_S9.
DR InterPro; IPR002469; Peptidase_S9B.

DR InterPro; IPR000379; Ser_estrs.
DR Pfam; PF00930; DPPIV_N; I.
DR Pfam; PF00326; Peptidase_S9; 1.
FT NON_TER 557 1
FT NON_TER 557 1
SQ SEQUENCE 557 AA; 62685 MW; 08AAA0944ACD0808 CRC64;
Alignment Scores:
Pred. No.: 1.55e-51 Length: 557
Score: 718.50 Matches: 185
Percent Similarity: 43.9% Conservative: 56
Best Local Similarity: 33.7% Mismatches: 159
Query Match: 24.3% Indels: 149
DB: 2 Gaps: 13
US-10-825-632-6 (1-1669) x QSTXJ2_ANOGA (1-557)
Qy 5 GGATACACCAAACTCTAAAGTCACCTTTTAAGATGTCAGAAATAATGATTGATGCTGAAGGA 64
Db 36 GlyThrProAenAlaLysSerLysLeuValGlnPheArgLeuSerGluAenLeu 55
Qy 65 AGGATCATAGATGTCATAGATAAGGAACCTAATTCAACCTTTTGAGATTCTATTTCGAAGGA 124
Db 56 ArgIleThrAspValCysIleLysGluLeuGlnCysProLeuThrPheAlaPheProTyr 75
Qy 125 GTTGAATATATGCCAGAGCTGGATGGACTCTCTGAGGGAATAATATGCTTGCTCATCTA 184
Db 76 LeuGluTyrIleValArgValGlyTyrThrProaspSerArgTyrValPalaGlnLeu 95
Qy 185 CTAGATCGCTCCAGACTCGCTCAGATAGTGTGTGATCTCACCTGAATATTATATCCCA 244
Db 96 LeuAspArgProGlnArgLeuGluLeuValLeuLeuProValaspAenPheCysSer 115
Qy 245 GTAGAGATGATGTTATGGAAAGGAGAGACTCATTGAGTCAGTGCCTGATCTCTGTGACG 304
Db 116 ThrAla-----AlaArgArgArgProThrAlaGlyArgProProGlyAlaMet 133
Qy 305 CCCTAATATTATCTAGAAACACACACATCTGGATAAATATCCATGACATCTTTCAT 364
Db 134 GlyProGlyIleTyrThrGluThrSerSerTyrValAenValHisAspValLeuGln 153
Qy 365 GTTTTCCCAAGTCACGAGAGAAATTCAGTTTATTTTGGCTCTGAAATGCAAAACA 424
Db 154 ---PheValGluLeuSerGluGlnGluVal-----Thr 163
Qy 425 GGTTCCTCATTTATACAAATTCATCTATTTTAAAGGAAGCAAAATAT-----AAA 478
Db 164 ThrAenSerHisProThrAlaHisProLysGlnArgGluGlyAlaHisGlyThrGlu 183
Qy 479 CGATCCAGTGGTGGCTGCCTGCTCCAGTGAATTTCAAGTGTCTCT----- 523
Db 184 ValSerSerGlyGlyAlaGlyAla---ThrAspHisSerLeuProSerMetAlaCysIle 202
Qy 524 -----ATCAAGAGAGATAGCAATTTACCAGTGTGTAATGGGAA 562
Db 203 GlySerThrLeuValAlaArgIleValGlnLysValThrLeuThrGlyGlyAspTyrGlu 222
Qy 563 GTTCTGGCCGCGCATGATCTAATATCCAAAGTTGATGAAGTCAGAGGCTGGTATATTTT 622
Db 223 ValLeuGlyArg-----AsnValTyrTyrAspArgValArgGlnLeuValTyrPhe 239
Qy 623 GAAGCACAAGAGACTCCCTTTAGAGCAATCACCTGTACGTAGTACGTAAATCTCT 682
Db 240 MetGlyLeuArgGluThrProLeuGluLysHisLeuTyrValValSerLeuAlaGlnPro 259
Qy 683 GGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACATTTCTTGTGTCATCAGTCAGAC 742
Db 260 AenGlnLeuArgLeuLeuThrMetProGlyTyrSerPheThrVal-----GluPhe 276
Qy 743 TGTGACTTCTTTT---ATAAGTAAGTATAGTAACAGAGAATCCACACTGTGTGCTCCCTT 799
Db 277 AenAspLeuValArgIleAlaHisAspSerAenThrAlaAenGlyAenGlyCysSer---- 295

Db 431 MetVal-----SerSerAspAspIleHisIleAspGluLysArgLysLeuValTyr 447
QY 620 TTTCGAAGCACCAGACTCCCTTTAGAGCATCACCCTGACGTAGTACGTAAAT 679
Db 448 PheThrGlyThrLysAspThrCysLeuGluGlnHisLeuTyrValThrArgPheAspLys 467
QY 680 CCT---GGAGAGGTGACAGGCTGACCTGCGGTCTACTACAT---TCTTGCTGCATC 733
Db 468 ProAsnSerGluIleLysArgLeuSerHisAlaAsnPheSerHisArgSerIleSerIle 487
QY 734 AGTCAGCAGCTGTGACTTCTTTATAGTAAGTATAGTAACCAAGAAAT----- 781
Db 488 SerSerAsnPheLysLysPheIleThrThrTyrSerAsnIleSerThrIleSerLysThr 507
QY 782 -----CCACAC 787
Db 508 GluValPheAspLeuIleTyrAsnAspAspAsnAsnAspAspAspIleTyrProIle 527
QY 788 TGTGTGCTCCCTTTACAGCTATCAAGTCTCCTGAAGTACCCCACTTGCAAAACAAGAA 847
Db 528 ValLysSerSerPhePheIleAsnAspAspAspAspAspAspAspAspAspLysLys 547
QY 848 TTTTGGGCCACCACTTTTGGATTTCAGCAGGTCTCTCTCTGACTACTCTCCACAGAATT 907
Db 548 IleAsnIleAsnIle-----ProLysIle 555
QY 908 TTCTCTTTTCAAACTACTACTGATTTACATTGATGGATGCTCTACAGCCTCATGAT 967
Db 556 PheAsnPheLysAsnSerLysGlyValThrIleTyrGlyGlnTyrThrLeuProSerAsp 575
QY 968 CTACAGCCTGGAAGAAATCTACTGCTGCTGCTCATATATGCTGCTCT----- 1018
Db 576 TyrSerLysAspLysIleProThrValValTyrValTyrGlyGlyProHisValGln 595
QY 1018 ----- 1018
Db 596 IleValArgAsnGlnTyrAsnTyrIleLysGlnHisTyrThrAsnPheGlyPheIleGln 615
QY 1018 ----- 1018
Db 616 ValMetIleAspAsnValGlySerAlaAsnArgGlyLeuGluPheGluSerHisIleArg 635
QY 1018 ----- 1018
Db 636 GluLysMetGlyGlnValGluIleGlyAspGlnValGluGlyIleAsnTyrLeuValGly 655
QY 1018 ----- 1018
Db 656 AsnAspIleValSerIleAspValAsnArgIleAlaIleSerGlyTrpSerTyrGlyGly 675
QY 1019 -----CAGGTTGCTATTGCTGGG 1036
Db 676 TyrAsnSerLeuMetAlaIleSerGlnArgProAspValPheLysIleAlaValCysGly 695
QY 1037 GCCCAGTCACTCTGGGATCTTCTATGATACAGGATACAGGAACGCTATATGGGTCAC 1096
Db 696 AlaProValSerAspTrpArgLeuTyrAsnThrGlyTyrThrGluArgTyrMetAsnVal 715
QY 1097 CCTGACCAGATGAACAGGCTATTACTAGGATCTGGCCATGCGCAAGCAGAAAGTTC 1156
Db 716 ProGlnAspAsnIleAspGlyTyrLysLeuGlyAspThrHisTyrSer-----Phe 733
QY 1157 CCTCTGAACCAATCGTTACTGCTTATCATGTTCTCTGGATGATGATGCTCAATTT 1216
Db 734 ProThrGluLysArgLeuLeuIleHisGlyLeuGlnAspGluAsnValHisPhe 753
QY 1217 GCATACACAGTATATTACTAGTATTTTGTAGGCGCTGGAAGCCATATGATTTACAG 1276
Db 754 SerAsnThrIleTyrIleIleAspHisLeuThrLysThrGlnLysProTyrIleLeuLys 773
QY 1277 ATCTATCTCAGAGACACAGCATATAAGATTCCTGAATCGGGAGAACATTATGAACGTG 1336
Db 774 ThrLeuProAsnGluArgHisGlyValArgAsnThrAspAsnArgIleTyrIleGlyLeu 793

QY 1337 CATCTTTTGCACTACCTTCAAGAAAAACCTT 1366
Db 794 PheValIleAsnHisLeuLysLeuLysAsnLeu 803
RESULT 15
Q965K3_CABEL
ID Q965K3_CABEL PRELIMINARY; PRT: 927 AA.
AC Q965K3;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Dipeptidyl peptidase four (IV) family protein 3, isoform b.
GN Name=dpf-3; ORFNames=K02F2.1;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Bristol N2;
RX MEDLINE=95069613; PubMed=9851916;
RG The C. elegans sequencing consortium;
RT "Genome sequence of the nematode C. elegans: a platform for
investigating biology";
RL Science 282:2012-2018 (1998).
DR EMBL AF043499; AAK84627.1; -, Genomic_DNA.
DR MEROPS; S09.016; -, Caenorhabditis elegans.
DR Ensembl; K02F2.1; Caenorhabditis elegans.
DR WormBase; WSGene0001056; dpf-3.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0003824; F:dipeptidyl-peptidase IV activity; IEA.
DR GO; GO:0004274; F:dipeptidyl-peptidase IV activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001375; Peptidase_S9.
DR InterPro; IPR002469; Peptidase_S98.
DR InterPro; IPR000379; Ser_estrs.
DR Pfam; PF00930; DPPIV_N; 1.
DR Pfam; PF00326; Peptidase_S9; 1.
DR Complete proteome.
SQ SEQUENCE 927 AA; 105366 MW; 6907C6AAEF829D0F CRC64;
Alignment Scores:
Pred. No.: 8,22e-30 Length: 927
Score: 462.00 Matches: 154
Percent Similarity: 38.7% Conservative: 79
Best Local Similarity: 25.6% Mismatches: 177
Query Match: 15.6% Indels: 192
DB: 2 Gaps:
US-10-825-632-6 (1-1669) x Q965K3_CABEL (1-927)
QY 5 GGTACAGCAATCCTTAAGTCACCTTTTAAGATGTCAGAAATATGATGCTGAAGGA 64
Db 351 GlyThrLysAsnAlaTyrSerThrLeuArgMet-----ValIleLeuGluAsnGly 367
QY 65 AGGATCATAGATGTCATAGATAAGAACTAATCAACCTTTTGAGATCTATTGAAGGA 124
Db 368 LysAlaTyrAspValProLeuLysAsp-----GluValIleTyrLysHis 382
QY 125 -----GTTGATATATTCCAGAGCTCGATGGAGCTCTCGAGGAAATATGCTTGG 175
Db 383 CysProPheTyrGluTyrIleThrArgAlaGlyPhePheSerAspGlyThrThrValTrp 402
QY 176 TCCATCCTACTAGTCTGCCAGACTCGCCTACAGATAGTGTGATC----- 223
Db 403 ValGlnValMetSerArgAspGlnAlaGlnCysSerLeuLeuLeuIleProTyrThrAsp 422
QY 224 -----TCACCT 229
Db 423 PheLeuLeuProGluLysGlySerIleLysGluAspAsnLeuLeuLeuSerThr 442

```
QY 230 GAATTATTTATCCAGTAGAAGATGATGTTATGAAAGGACAGAGACTCATTTGAGTCAGTG 289
Db 443 ApeLeuAsnMetGlyValTrpAspAap---LysSerHisGluGluThrMetGluLysPro 461
QY 290 CTTGATCTTGTCAGC-----CCACTAATTTCTATGAGAAACACACAGACATCTCGATA 343
Db 462 ProArgGlyLysLeuArgGlyThrValGlnIleHisLysAlaArgAsnAspTyrTrpIle 481
QY 344 AATATCCATGACATCTTTTCATGTTTCCCAAGTCCACGAAGAGAA-----ATTGAG 397
Db 482 AsnThrHisAsnAlaIleTyrProLeuLysIleThrAspGluGluHisProMetTyrGlu 501
QY 398 TTTATTTTTCCTCTGAATGCAAAACAGGTTTCCTCATTTATACAAAATTTACATCTATT 457
Db 502 PheIleTyrCysLeuGluLysProAsnGly--SerCysLeuAlaLeuIleSerAlaGlu 520
QY 458 TTAAGGAAGCAATATTAACGATCCAGTGGTGGCTGCTCCCAAGTGATTTCAAG 517
Db 521 LeuAspGlnAsnGlyTyr----- 526
QY 518 TGTCCCTATCAAGAGGAGATGCAATTACCACTGGTGAATGGGAAGTTCTTGGCCGCAT 577
Db 527 CysArgHisThrGluGluLysLeuLeuMetAlaGluAsnPheSerIle-----AsnLys 544
QY 578 GATCTAATATCAAGTTGATGAAGTCAGAAGCTGGTATATTTTGAAGGCACCAAGAC 637
Db 545 SerMetGlyIleValValAspGluValArgGluLeuValTyrTyrValAlaAsnGluSer 564
QY 638 TCCCTTTTGAAGTACACCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 697
Db 565 HisProThrGluTrpAsnIle---CysValSerHisTyrArgThrGlyGlnHisAlaGln 583
QY 698 CTGACTGACCGTGGCTACTCACATCTTCTGTCATGTCAGTCAGTCAGTCAGTCAGTCAG 757
Db 584 LeuThrGluSerGlyIle-----Cys-----PheLys 592
QY 758 AGTAAGTATAGTAACACAGAG-----AATCCACACTGTGTGTCCTTTTACAAAG 805
Db 593 SerGluArgAlaAsnGlyLysLeuAlaLeuAspLeuAspHisGlyPheAlaCysTyrMet 612
QY 806 CTATCA-----AGTCCTGAAGATGACCACTTGCACAAACAAAGGAATTT---TGG--- 853
Db 613 ThrSerValGlySerProAlaGlu-----CysArgPheTyrSerPheArgTrpLys 629
QY 854 -----GCCACCATTTTGGATTTCAGCAGGTCCT 880
Db 630 GluAsnGluValLeuProSerThrValTyrAlaAlaAsnIleThrValSerGlyHisPro 649
QY 881 ---CTTCTGAC-----TATACTCTCCAGAAATTTTCTCTTTTGAAGT---ACTACT 928
Db 650 GlyGlnProAspLeuHisPheAspSerProGluMetIleGluPheGlnSerLysLysThr 669
QY 929 GGATTTACATGTTGGATGCTCTACAGCCCTCATGATCTACAGCCCTGGGAAAGAAATAT 988
Db 670 GlyLeuMetHisTyrAlaMetIleLeuArgProSerAsnPheAspProTyrLysLysTyr 689
QY 989 CCTACTGCTGCTCATATATGTTGCTCT----- 1018
Db 690 ProValPheHisTyrValTyrGlyGlyProGlyIleGlnIleValHisAsnAspPheSer 709
QY 1018 ----- 1018
Db 710 TrpIleGlnTyrIleArgPheCysArgLeuGlyTyrValValValPheIleAspAsnArg 729
QY 1018 ----- 1018
Db 730 GlySerAlaHisArgGlyIleGluPheGluArgHisIleHisLysLysMetGlyThrVal 749
QY 1018 ----- 1018
Db 750 GluValGluAspGlnValGluGlyLeuGlnMetLeuAlaGluArgThrGlyGlyPheMet 769
QY 1018 ----- 1018

Db 770 AspMetSerArgValValValHisGlyTrpSerTyrGlyGlyTyrMetAlaLeuGlnMet 789
QY 1019 -----CAGGTTGCTATTGCTGGGCCCCAGTCACCTCTCTGG 1054
Db 790 IleAlaLysHisProAsnIleTyrArgAlaAlaIleAlaGlyGlyAlaValSerAspTrp 809
QY 1055 ATCTTCTATGATACAGGATACACGAAAGCTTATATGGGTCAACCTGACCAAGATGAACAG 1114
Db 810 ArgLeuTyrAspThrAlaTyrThrGluArgTyrMetGlyTyrPro---LeuGluGluHis 828
QY 1115 GGTATTACTTAGGATCTGCGCCATGCAACGACGAAAGTTCCCTCTGACCAACAAATCGT 1174
Db 829 ValTyrGlyAlaSerSerIleThrGlyLeuValGluLysLeuProAspGluProAsnArg 848
QY 1175 TTACTGCTCTTACATGTTTCTGATGAGATGTCCATTTTGCACATACACAGTATATTA 1234
Db 849 LeuMetLeuValHisGlyLeuMetAspGluAsnValHisPheAlaHisLeuThrHisLeu 868
QY 1235 CTGAGTTTTTTAGTAGGCTGGAAGCCATATGATTTTACAGATCTATCTCTCAGGAGAGA 1294
Db 869 ValAspGluCysIleLysLysGlyLysTrpHisGluLeuValIlePheProAsnGluArg 888
QY 1295 CACAGCATAGAGTTCTGAAATCGGAGAGAACATTATGAACTGATCTTTTGCACCTACCTT 1354
Db 889 HisGlyValArgAsnAsnAspAlaSerIleTyrLeuAspAlaArgMetMetTyrPheAla 908
QY 1355 CAAGAA 1360
Db 909 GlnGln 910

Search completed: May 2, 2006, 01:33:18
Job time : 424.39 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: May 2, 2006, 01:34:19 ; Search time 10.2279 Seconds
(without alignments)
2698.234 Million cell updates/sec

File: US-10-825-632-6
Perfect score: 2957
Sequence: 1 aaacagtcagcaaatccta.....aaaaaaaaaaaaaaaaaaaaa 1669

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 572060 seqs, 82675679 residues
Total number of hits satisfying chosen parameters: 1144120

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ n2p.model -DEV=xlp
-Q=/abss/ABSSWEB_spool/US10825632/runat_01052006_105950_3311/app_query.fasta_1
-DB=Issued Patents AA -QFMT=fastan -SUFFIX=n2p.ra1 -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=ptc -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-HOST=abss02p -USER=US10825632 @CGN 1.1.131 @runat_01052006_105950_3311
-NCFU=6 -ICFU=3 -NO MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents AA: *
1: /cgn2_6/prodata/1/iaa/5 COMB pep: *
2: /cgn2_6/prodata/1/iaa/6 COMB pep: *
3: /cgn2_6/prodata/1/iaa/H COMB pep: *
4: /cgn2_6/prodata/1/iaa/PCTUS COMB pep: *
5: /cgn2_6/prodata/1/iaa/RE COMB pep: *
6: /cgn2_6/prodata/1/iaa/backfiles1.pcp: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2482	83.9	465	2	US-10-070-464-5
2	2422	81.9	882	2	US-09-976-674-1
3	2422	81.9	882	2	US-10-070-464-1
4	1957.5	66.2	690	2	US-09-976-674-7
5	1808	61.1	658	2	US-09-976-674-19
6	1808	61.1	661	2	US-09-976-674-11
7	1540	52.1	613	2	US-09-976-674-21
8	1494.5	50.5	863	2	US-09-976-674-3
9	1494.5	50.5	882	2	US-09-976-674-23
10	1494.5	50.5	892	2	US-09-976-674-27
11	1445	48.9	879	2	US-09-976-674-33
12	1445	48.9	879	2	US-09-976-674-35

13	1377.5	46.6	310	2	US-09-794-236-4	Sequence 4, Appli
14	1377.5	46.6	310	2	US-10-070-464-3	Sequence 3, Appli
15	1094.5	37.0	832	2	US-09-976-674-29	Sequence 23, Appl
16	1094.5	37.0	832	2	US-09-976-674-31	Sequence 31, Appl
17	1045	35.3	819	2	US-09-976-674-37	Sequence 37, Appl
18	1045	35.3	819	2	US-09-976-674-39	Sequence 39, Appl
19	860.5	29.1	360	2	US-10-070-464-7	Sequence 7, Appli
20	471.5	15.9	518	2	US-09-976-674-25	Sequence 25, Appl
21	309.5	10.5	771	2	US-09-462-284-2	Sequence 2, Appli
22	309.5	10.5	771	2	US-09-079-592-2	Sequence 2, Appli
23	305	10.3	676	2	US-09-518-550-42	Sequence 42, Appl
24	305	10.3	723	2	US-09-518-550-29	Sequence 29, Appl
25	296.5	10.0	710	2	US-09-518-550-28	Sequence 28, Appl
26	294	9.7	102	2	US-09-270-767-44517	Sequence 44517, A
27	288	9.7	732	2	US-09-518-550-30	Sequence 30, Appl
28	281.5	9.5	981	2	US-09-902-540-16812	Sequence 16812, A
29	277	9.4	818	2	US-09-462-845-3	Sequence 3, Appli
30	277	9.4	818	2	US-10-402-312-3	Sequence 3, Appli
31	277	9.4	818	2	US-10-401-437-3	Sequence 3, Appli
32	277	9.4	818	2	US-10-402-067-3	Sequence 3, Appli
33	277	9.4	818	2	US-10-401-436-3	Sequence 3, Appli
34	267.5	9.0	931	2	US-09-079-592-11	Sequence 11, Appl
35	263	8.9	54	2	US-09-621-976-5647	Sequence 5647, Ap
36	249.5	8.4	706	2	US-09-976-674-41	Sequence 41, Appl
37	249.5	8.4	743	2	US-10-363-937-4	Sequence 4, Appli
38	249.5	8.4	796	2	US-09-976-674-5	Sequence 5, Appli
39	244	8.3	737	2	US-09-902-540-11421	Sequence 11421, A
40	238	8.0	766	2	US-10-002-593-6	Sequence 6, Appli
41	238	8.0	766	2	US-09-949-016-6146	Sequence 6146, Ap
42	238	8.0	766	2	US-10-423-714-6	Sequence 6, Appli
43	234	7.9	755	4	PCT-US93-07923-3	Sequence 3, Appli
44	234	7.9	759	4	PCT-US93-07923-2	Sequence 2, Appli
45	234	7.9	766	1	US-08-230-491A-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-10-070-464-5
; Sequence 5, Application US/10070464
; Patent No. 6881564
; GENERAL INFORMATION:
; APPLICANT: ABBOTT, Catherine Anne
; APPLICANT: GORRELL, Mark Douglas
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES
; FILE REFERENCE: GH-007
; CURRENT APPLICATION NUMBER: US/10/070,464
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: PCT/AU00/01085
; PRIOR FILING DATE: 2000-09-11
; PRIOR APPLICATION NUMBER: AU PQ5709
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: AU PQ2762
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 465
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-070-464-5

Alignment Scores:
Pred. No.: 3.76e-283 Length: 465
Score: 2482.00 Matches: 465
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 83.9% Indels: 0
DB: 2 Gaps: 0

US-10-825-632-6 (1-1669) x US-10-070-464-5 (1-465)

QY 2 ACAGGTACAGCAAAATCTTAAGTCACCTTTTAAGATGTCAGAAATATGATTGATCTGAA 61

Db 1 ThrGlyThrAlaAenProLysValThrPheLysMetSerGluLeuMetIleAspAlaGlu 20
QY 62 GGAAGGATCATAGTGTATAGTAGAGAACTAATTCACCTTTTGTGATCTATTGAA 121
Db 21 GlyArgIleIleAspValIleAspLysGluLeuIleGlnProPheGluLeuPheGlu 40
QY 122 GGAATTGAATATATCCAGAGCTGGATGGACTCTCTGAGGAAAATATGCTTGTCATC 181
Db 41 GlyValGluTyrIleAlaArgAlaGlyIrpThrProGluGlyLysTyrAlaIrpSerIle 60
QY 182 CTACTAGATCGCTCCAGACTGCCCTACAGATAGTGTGATCTCACTGGAATATTATTC 241
Db 61 LeuLeuAspArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIle 80
QY 242 CCAGTAGAGATGATGTTATGGAAGGCGAGAGACTCATTTGAGTCAGTGCCTGATTCGTG 301
Db 81 ProValGluAspValMetGluArgGlnArgLeuIleGluSerValProAspSerVal 100
QY 302 AC6CCACTAATTATCATGAAGAAACAACAGACATCTGGATAAATATCCATGACATCTTT 361
Db 101 ThrProLeuIleIleTyrGluGluThrThrAspIleTrpIleAenIleHisAspIlePhe 120
QY 362 CATGTTTTTCCCAAGTCACGAAGAGAAATTTGAGTTTATTTTGGCTCTGATGCAAA 421
Db 121 HisValPheProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLys 140
QY 422 ACAGGTTCCGTCATTTATACAAATTTACATCTATTTTAAAGGAAAGCAATATAACGA 481
Db 141 ThrGlyPheArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArg 160
QY 482 TCCAGTGGTGGCTGCTCCCAAGTCATTTCAAGTCTCCTATCAAGAGGAGATAGCA 541
Db 161 SerSerGlyGlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAla 180
QY 542 ATTACAGTGGTGAATGGAGTCTTCCGCGGATGATGATCAATATCCAGTTGATGAA 601
Db 181 IleThrSerGlyGluTyrGluValLeuGlyArgHisGlySerAenIleGlnValAspGlu 200
QY 602 GTCAGAGGCTGGTATTTTGAAGGACCAAGACTCCCTTTAGAGCATCCTGTATC 661
Db 201 ValArgAspLeuValTyrPheGluGlyThrLysAspSerProLeuGluHisIleTyr 220
QY 662 GTAGTCAGTTAGTAAATCCTGGAGAGGTGACAAAGGTGACTGACCGTGCTACTACAT 721
Db 221 ValValSerTyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHis 240
QY 722 TCTTGTGTCATCAGTCAGACTGTGACTCTTTTATAAGTAAATATAGTAACAGAGAAAT 781
Db 241 SerCysCysIleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAen 260
QY 782 CCACACTGTGTGCTCCCTTTACAAGCTATCAAGTCTCGAAGTACCCCACTTGCAAAACA 841
Db 261 ProHisCysValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThr 280
QY 842 AAGGAATTTTGGGCCACCATTTTGGATTCAGCAGGTCTCTTCTGACTATATCTCTCCA 901
Db 281 LysGluPheTrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProPro 300
QY 902 GAAATTTCTCTTTTGAAGTACTACTGGATTTTACATTTGATGGGATGCTCTACAGCT 961
Db 301 GluIlePheSerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysPro 320
QY 962 CATGATCTACAGCTCGGAAAGAAATATCCTACTGTGCTGTTCATATATGCTGCTCAG 1021
Db 321 HisAspLeuGlnProGlyLysLeuTyrProThrValLeuPheIleTyrGlyLysProGln 340
QY 1022 GTTGTATTTGCTGGGCCCGAGTCACCTGTGGATCTTCTATGATACAGATACAGGAA 1081
Db 341 ValAlaIleAlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGlu 360
QY 1082 GCTTATATGGTACCCTTGACCAAGTAAACAGGGCTATTTACTTAGATCTCTGGCCATG 1141

Db 361 ArgTyrMetGlyHisProAspGlnAenGluGlnGlyTyrTyrLeuGlySerValAlaMet 380
QY 1142 CAAGCAGAAAGTTCCCTCTCTGAACCAAAATCGTTTACTGCTCTTACATGTTTCTCGAT 1201
Db 381 GlnAlaGluLysPheProSerGluProAenArgLeuLeuLeuHisGlyPheLeuAsp 400
QY 1202 GAGAAATCCATTTTGCACATACCAAGTATATTTACTGAGTTTTTTAGTGAGGCTGGAAG 1261
Db 401 GluAenValHisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLys 420
QY 1262 CCATATGATTTACAGATCTATCTCTCAGGAGACACAGACATAAGAGTTTCTCGAATCGGA 1321
Db 421 ProTyrAspLeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGly 440
QY 1322 GAACATTATCAACTGCATCTTTTGCATACCTTCAAGAAACCTTGGATCACGATTGCT 1381
Db 441 GluHisTyrGluLeuHisLeuLeuHisTyrLeuGlnGluAenLeuGlySerArgIleAla 460
QY 1382 GCTCTAAAGTGATA 1396
Db 461 AlaLeuLysValIle 465
RESULT 2
US-09-976-674-1
; Sequence 1, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akisanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junich, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 882
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-1
Alignment Scores:
Pred. No.: 6,95e-276 Length: 882
Score: 2422.00 Matches: 465
Percent Similarity: 82.3% Conservative: 0
Best Local Similarity: 82.3% Mismatches: 0
Query Match: 81.9% Indels: 100
DB: 2 Gaps: 1
US-10-825-632-6 (1-1669) x US-09-976-674-1 (1-882)
QY 2 ACAGGTACAGCAAAATCCCTAAAGTCACCTTTAAGATGTCAGAAATATGATGATGCTGAA 61
Db 318 ThrGlyThrAlaAenProLysValThrPheLysMetSerGluLeuMetIleAspAlaGlu 337
QY 62 GGAAGATCATAGTGTCTATAGTAAGCAACTAATTCACCTTTGAGATCTATTGAA 121
Db 338 GlyArgIleIleAspValIleAspLysGluLeuIleGlnProPheGluLeuPheGlu 357
QY 122 GGAATTGAATATATTGCCAGAGCTGGATGAGTCTCCTCAGGAGAAAATATGCTTGCTCATC 181
Db 358 GlyValGluTyrIleAlaArgAlaGlyIrpThrProGluGlyLysTyrAlaIrpSerIle 377
QY 182 CTACTAGATCGCTCCAGACTCGGCTACAGATAGTGTGATCTCACTGAAATATTATTC 241
Db 378 LeuLeuAspArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIle 397
QY 242 CCAGTAGAGATGATGTTATGGAAGGCGAGAGACTCATTCAGTCAGTGCCTGATTCGTG 301

```
Db      398 PtoValGluAspValMetGluArgGlnArgLeuIleGluSerValProAspSerVal 417
QY      302 A6CCCACTAATTATCTATGAGAAACAACAGACATCTGGATAAATATCCATGACATCTTT 361
Db      418 ThrProLeuIleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePhe 437
QY      362 CATGTTTTCCTCCCAAGTCACGAGAGAAATTCAGTTTATTTTTCCTCTCGAATCAAA 421
Db      438 HisValPheProGlnSerHisGluGluIleGluPheIlePheAlaSerGluCysLys 457
QY      422 ACAGGTTTCCTCGTCATTATACAAAATTACATCTATTTTAAAGGAAAGCAAAATATAACGA 481
Db      458 ThrGlyPheArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArg 477
QY      482 TCCAGTGTGGGCTGCTCTCCAAAGTGATTTCGAAGTGTCTCTATCAAAAGAGGAGATAGCA 541
Db      478 SerSerGlyGlyLeuProAlaProSerAspPheLysCysProIleLysGluIleAla 497
QY      542 ATTACAGTGTGAATGGGAAGTCTTGGCCGGCATGGATCTAATATCCAAAGTTGATGAA 601
Db      498 IleThrSerGlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGlu 517
QY      602 GTCAGAGGCTGGTATATTTCAAGGCACCAAGCTCCCTTTAGAGCATCACCTGTAC 661
Db      518 ValArgArgLeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeuTyr 537
QY      662 GTAGTCAGTACGTAACTCTCGAGAGGTGACAGGCTGACTGACCGTGGCTACTCACAT 721
Db      538 ValValSerTyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHis 557
QY      722 TCTGTGTCATCAGTCAGTCATGTCATCTCTTTTATAAGTAAGTAGTAGTAACCAAGAAGAT 781
Db      558 SerCysCysIleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsn 577
QY      782 CCACTACTGTGTGCTCTTTACAAAGCTATCAAGTCTCTGAAGTAGCCCACTTGCAGAAACA 841
Db      578 ProHisCysValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThr 597
QY      842 AAGGAATTTGGCCACCACTTTTGATTCAGCAGGTCTCTCTCTGACTATATCTCTCCA 901
Db      598 LysGluPheTrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProPro 617
QY      902 GAATTTCTCTTTTGAAGTACTACTCGATTACATTGATGGATGGCTCTCAAGCCT 961
Db      618 GluIlePheSerPheGlnSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysPro 637
QY      962 CATGATCTACAGCTCGAAGAAATATCTCTACTGTCTGTTCATATATGGTGGTCTCTCAG 1021
Db      638 HisAspLeuGlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGln 657
QY      1021 ----- 1021
Db      658 ValGlnLeuValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAla 677
QY      1021 ----- 1021
Db      678 SerLeuGlyTyrValValValIleAspAsnArgLysSerCysHisArgLysLeuLys 697
QY      1021 ----- 1021
Db      698 PheGluGlyAlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGly 717
QY      1021 ----- 1021
Db      718 LeuGlnTyrLeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGly 737
QY      1021 ----- 1021
Db      738 TrpSerTyrGlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArg 757
QY      1022 GTTGCTATTGCTGGGCCCCAGTCATCTGTGGATCTCTGTGATGATCATAGATACAGGATACACGAA 1081
Db      758 ValAlaIleAlaGlyAlaProValThrLeuTyrPheThrAspThrGlyTyrThrGlu 777
```

```
QY      1082 CGTTATATGGTCACCTGACCAGAAATGAACAGGCTATTACTTTAGGATCTGTGGCCATG 1141
Db      778 ArgTyrMetGlyHisProAspGlnAsnGlnGlnGlyTyrTyrLeuGlySerValAlaMet 797
QY      1142 CAAGCAGAAAAGTTCCTCTCTGACCAAAATCGTTTACTGCTCTTACATGGTTTCTCGAT 1201
Db      798 GlnAlaGluLysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAsp 817
QY      1202 GAGAAATGTCATTTCACATACCAAGTATATTAATCTAGTGTGTTTGTAGTGGCTGGAAAG 1261
Db      818 GluAsnValHisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLys 837
QY      1262 CCATATGATTACAGATCTCTCTCAGGAGACACACATAAGAGTTCTCTGAATCGGA 1321
Db      838 ProTyrAspLeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGly 857
QY      1322 GAACATTATGAACCTGCATCTTTTGCATCTCTTCAAGAAAAACCTTGGATCACGATTGCT 1381
Db      858 GluHisTyrGluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAla 877
QY      1382 GCTCTAAAAAGTGATA 1396
Db      878 AlaLeuLysValIle 882

RESULT 3
US-10-070-464-1
; Sequence 1, Application US/10070464
; Patent No. 6881564
; GENERAL INFORMATION:
; APPLICANT: ABBOTT, Catherine Anne
; APPLICANT: GORRELL, Mark Douglas
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES
; FILE REFERENCE: GH-007
; CURRENT APPLICATION NUMBER: US/10/070,464
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: PCT/AU00/01085
; PRIOR FILING DATE: 2000-09-11
; PRIOR APPLICATION NUMBER: AU PQ5709
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: AU PQ2762
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 882
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-070-464-1

Alignment Scores:
Pred. No.: 6,95e-276 Length: 882
Score: 2422.00 Matches: 465
Percent Similarity: 82.3% Conservative: 0
Best Local Similarity: 82.3% Mismatches: 0
Query Match: 81.9% Indels: 100
DB: 2 Gaps: 1

US-10-825-632-6 (1-1669) x US-10-070-464-1 (1-882)
QY      2 ACAGGTACAGCAAAATCTCTAAAGTCACCTTTTAAGATGTCAGAAATAATGATGCTGAA 61
Db      318 ThrGlyThrAlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGlu 337
QY      62 GGAAGGATCATAGATGTCATAGATAAGGAACCTAACTTTTCAACCTTTTTCAGATTCTATTGAA 121
Db      338 GlyArgIleIleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGlu 357
QY      122 GGAGTTCAATATATTTGCCAGAGCTGGATGGACTCTCTCGAGGGAAAAATATCTTGTGTCATC 181
Db      358 GlyValGluTyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIle 377
QY      182 CTACTAGATCGCTCCAGACTCGCCTACAGATAGTGTGTTGATCTCACCTGAATATTATTATC 241
```

Db 378 LeuLeuAspArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIle 397
QY 242 CAGTAGAAGCATGCTATGAAAGGAGAGACCTATTGAGTCAGTGGCTGATCTGCG 301
Db 398 ProValGluAspValMetGluArgGlnArgLeuIleGluSerValProAspSerVal 417
QY 302 ACCCCACTAATATCTATCAAGAAACACAGCATCTGGATAAATATCCATGACATCTTT 361
Db 418 ThrProLeuIleIleTyGluGlnThrThrAspIleIlePheAsnIleHisAspIlePhe 437
QY 362 CATGCTTTTCCCAAGTCACAGAGAAATTTGAGTTATTTTGGCTCTGAAATGCAA 421
Db 438 HisValPheProGlnSerHisGluGluIleGluPheIlePheAlaSerGluCysLys 457
QY 422 ACAGGTTTCGTCATTTATACAAATTTACATCTATTTTAAAGGAAAGCAAAATATAACGA 481
Db 458 ThrGlyPheArgHisLeuTyrlsIleThrSerIleLeuLysGluSerLysTyrlsArg 477
QY 482 TCCAGTGGTGGCTGCTGCTCCAAAGTGATTTCAAGTGCTCTATCAAGAGGAGATAGCA 541
Db 478 SerSerGlyGlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAla 497
QY 542 ATTACAGTGGTGAATGGGAAGTCTTGGCCGGCATGGATCTAATATCCAAAGTTGATGAA 601
Db 498 IleThrSerGlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGlu 517
QY 602 GTCAGAGGCTGGTATATTTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTAC 661
Db 518 ValArgArgLeuValTyrlsPheGluGlyThrLysAspSerProLeuGluHisLeuTyrls 537
QY 662 GTAGTCAGTACCTAAATCTGGAGAGGTGACAGGCTGACTCACCGTGGCTACTACAT 721
Db 538 ValValSerTyrlsValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrlsHis 557
QY 722 TCTTGTGTCATCAGTCACACTGTGACTTCTTATAGTAAGTATAGTAAACAGAGAGAT 781
Db 558 SerCysCysIleSerGlnHisCysAspPhePheIleSerLysTyrlsSerAsnGlnLysAsn 577
QY 782 CCACACTGTGCTCCCTTTACAAGCTATCAAGCTGAGCTGAGATCAACCACTTGCAGAA 841
Db 578 ProHisCysValSerLeuTyrlsLeuSerSerProGluAspProThrCysLysThr 597
QY 842 RAGGAATTTGGCCGACCATTTTGGATTCAGAGTCTCTCTCTGACTATATCTCTCCA 901
Db 598 LysGluPheTrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrlsProPro 617
QY 902 GAAATTTCTCTTTGAAAGTACTGGAATTTACATGATGGGATGCTCTACAAGCCT 961
Db 618 GluIlePheSerPheGluSerThrThrGlyPheThrLeuTyrlsGlyMetLeuTyrlsPro 637
QY 962 CATGATCTACAGCTCGGAAAGAAATATCTACTGTGCTGTTCATATATGTTGGTGGTCTCAG 1021
Db 638 HisAspLeuGlnProGlyLysTyrlsProThrValLeuPheIleTyrlsGlyProGln 657
QY 1021 ----- 1021
Db 658 ValGlnLeuValAsnAsnArgPheLysGlyValLysTyrlsPheArgLeuAsnThrLeuAla 677
QY 1021 ----- 1021
Db 678 SerLeuGlyTyrlsValValValValIleAspAsnArgGlySerCysHisArgGlyLeuLys 697
QY 1021 ----- 1021
Db 698 PheGluGlyAlaPheLysTyrlsMetGlyGlnIleGluIleAspAspGlnValGluGly 717
QY 1021 ----- 1021
Db 718 LeuGlnTyrlsLeuAlaSerArgTyrlsAspPheIleAspLeuAspArgValGlyIleHisGly 737
QY 1021 ----- 1021

Db 738 TrpSerTyrlsGlyGlyTyrlsLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArg 757
QY 1022 GTTCTATTCTCTGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAA 1081
Db 758 ValAlaIleAlaGlyAlaProValThrLeuTrpIlePheTyrlsAspThrGlyTyrlsGlu 777
QY 1082 CGTTATATGGGTCAACCTTCCAGCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCCATG 1141
Db 778 ArgTyrlsMetGlyHisProAspGlnAsnGluGlnGlyTyrlsLeuGlySerValAlaMet 797
QY 1142 CAAGCAGAAAGTTCCCTCTCGAACCAATCGTTTACTGCTCTTACATGGTTTCTCTGGAT 1201
Db 798 GlnAlaGluLysPheProSerGluProAsnArgLeuLeuLeuLeuHisGlyPheLeuAsp 817
QY 1202 GAGAAATGTCATATTTCCACATACCAAGTATATTACTAGTGTATTTTAGTGAGGCTGGAAG 1261
Db 818 GluAsnValHisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLys 837
QY 1262 CCATATGATTTACAGATCTATCTCAGGAGAGACACAGATAGAGTTTCTTGAATCGGA 1321
Db 838 ProTyrlsAspLeuGlnIleTyrlsProGlnArgHisSerIleArgValProGluSerGly 857
QY 1322 GAACATTATGAATGCTCATCTTTTGGCTACCTTCAAGAAACCTTGGATCAGCTATTGCT 1381
Db 858 GluHisTyrlsGluLeuHisLeuLeuHisTyrlsLeuGlnLeuLeuGlySerArgIleAla 877
QY 1382 GCTCTAAAGTGTATA 1396
Db 878 AlaLeuLysValIle 882
RESULT 4
US-09-976-674-7
; Sequence 7, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976.674
; PRIOR FILING DATE: 2001-10-12
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 690
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-7
Alignment Scores:
Pred. No.: 3, 61e-221 Length: 690
Score: 1957.50 Matches: 372
Percent Similarity: 99.5% Conservative: 1
Best Local Similarity: 99.2% Mismatches: 0
Query Match: 66.2% Indels: 2
DB: 2 Gaps: 1
US-10-825-632-6 (1-1669) x US-09-976-674-7 (1-690)
QY 2 ACAGTACAGCAATCTTAAGTCACTTTAAGATGTCAGAAATATGATTGATGCTGAA 61
Db 318 ThrGlyThrAlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGlu 337
QY 62 GGAAGATCATAGATGTCATAGATAGCAACTAATTAACCTTTTCTGAGATTTCTATTGAA 121
Db 338 GlyArgIleIleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGlu 357
QY 122 GGAGTTGAATATATTGCCAGCTGGATGCTCTCCTGAGGAAATATATCTTGGTCCATC 181

```
Db 358 GlyValGluTyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIle 377
QY 182 CTACTAGATCGCTCCAGACTCGCTTACAGATGAGTGTGATCTCACCTGAATATTATTC 241
Db 378 LeuLeuAspArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIle 397
QY 242 CCAGTGAAGATGATGTTAAGAACGACAGACTCATTGACTCAGTCCCTGATCTCTG 301
Db 398 ProValGluAspValMetGluArgGlnArgLeuIleGluSerValProAspSerVal 417
QY 302 ACCCACTAATATCTATGAAGAACACACAGACTCTGATAAATATCCATGACATCTTT 361
Db 418 ThrProLeuIleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePhe 437
QY 362 CATGTTTTTCCCAAGTCACGAGAGGAAATGAGTTATATTTTGGCTCTGAATGAAA 421
Db 438 HisValPheProGlnSerHisGluGluIleGluPheIlePheAlaSerGluCysLys 457
QY 422 ACAGTTTCCGTCATTTATACAAATTAATCATCTATTTTAAAGGAACCAATATAACGA 481
Db 458 ThrGlyPheArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArg 477
QY 482 TCACAGTGGTGGCTCGCTGCTCAAGTGTCTTCAAGTGTCTTCAAGAGAGAGATAGCA 541
Db 478 SerSerGlyGlyLeuProAlaProSerAspPheLysCysProIleLysGluIleAla 497
QY 542 ATTACCACTGGTGAATGGGAAAGTCTTCTGCGGATGGATCTTAATATCCAAAGTTGATA 601
Db 498 IleThrSerGlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGlu 517
QY 602 GTGAGAAGGCTGGTATATTTTGAAGGCACCAAGACTCCCTTTAGACATCACCTGTAC 661
Db 518 ValArgArgLeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeuTyr 537
QY 662 GTAGTCAGTACGTAATCTGGAGAGGTGACAAAGCTCAGTACCGTGGTACTTCACAT 721
Db 538 ValValSerTyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHis 557
QY 722 TCTTGCTGTCATCAGCTAGCTGTGACTTCTTTTATAGTAAGTATAGTAACAGAGAAAT 781
Db 558 SerCysCysIleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsn 577
QY 782 CCACTGTGTGCTCCCTTTACAGCTATCAAGTCTTGAAGTACCCCACTTGCAAAACA 841
Db 578 ProHisCysValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThr 597
QY 842 AAGGAATTTTGGGCCACCATTTTGGATTTCAGAGGTCTCTTCTGACTATATCTCTCCA 901
Db 598 LysGluPheTrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProPro 617
QY 902 GAAATTTTCTCTTTTGAAGTACTACTGGATTACATGTTATGGGATGCTCTACAGCT 961
Db 618 GluIlePheSerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysPro 637
QY 962 CATGATCTACAGCTGGGAAGAAATATCTACTGCTGCTGTTCATATATGCTGCTCTCAG 1021
Db 638 HisAspLeuGlnProGlyLysTyrProThrValLeuPheIleTyrGlyLysArg 656
QY 1022 GTTGTATTGCTGGGGCCCCAGTCACTGTGATCTTCTATGATACAGGATACAGGAA 1081
Db 657 -LeuLeuLeuLeuGlyProGlnSerLeuCysGlySerSerMetIleGlnAspThrArgAs 676
QY 1082 CDTTATATGGTGCACCTGACCAAGTGAACAGGGCTATTACT 1124
Db 676 nValIleTrpValThrLeuThrArgMetAsnArgAlaIleThr 690
```

```
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19
; LENGTH: 658
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-19

Alignment Scores:
Pred. No.: 1,5e-203 Length: 658
Score: 1808.00 Matches: 338
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 61.1% Indels: 0
DB: 2 Gaps: 0

US-10-825-632-6 (1-1669) x US-09-976-674-19 (1-658)
QY 2 ACAGGTACAGCAAAATCTTAAAGTCACCTTTTAAGATGTCAGAAATATGATGCTGAA 61
Db 318 ThrGlyThrAlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGlu 337
QY 62 GGAAGGATCATPAGATGTCATAGATAAGAACTAATTAACCTTTTGAGATCTTATTGAA 121
Db 338 GlyArgIleIleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGlu 357
QY 122 GGAGTTGAATATATGTCAGAGCTGGAGTGGACTCTCTGAGGAAATATCTTGGTCCATC 181
Db 358 GlyValGluTyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIle 377
QY 182 CTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCCTCAATATTATTTATC 241
Db 378 LeuLeuAspArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIle 397
QY 242 CCAGTGAAGATGATGTTATGGAAGCAGAGACTCATTTGAGTCAGTCCCTGATCTGTG 301
Db 398 ProValGluAspValMetGluArgGlnArgLeuIleGluSerValProAspSerVal 417
QY 302 ACACCACTAATATCTATGAGAACAGACATCTGGATAAATATCCATGACATCTTT 361
Db 418 ThrProLeuIleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePhe 437
QY 362 CATGTTTTTCCCAAGTCACGAGAGGAAATGAGTTATATTTTGGCTCTGAATGAAA 421
Db 438 HisValPheProGlnSerHisGluGluIleGluPheIlePheAlaSerGluCysLys 457
QY 422 ACAGTTTCCGTCATTTATACAAATTAATCATCTATTTTAAAGGAACCAATATAACGA 481
Db 458 ThrGlyPheArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArg 477
QY 482 TCCAGTGGTGGCTCGCTGCTCCAAGTGTCTTCAAGTGTCTTCAAGAGAGAGATAGCA 541
Db 478 SerSerGlyGlyLeuProAlaProSerAspPheLysCysProIleLysGluIleAla 497
QY 542 ATTACCACTGGTGAATGGGAAAGTCTTCTGCGGATGGATCTTAATATCCAAAGTTGATA 601
Db 498 IleThrSerGlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGlu 517
QY 602 GTGAGAAGGCTGGTATATTTTGAAGGCACCAAGACTCCCTTTAGACATCACCTGTAC 661
Db 518 ValArgArgLeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeuTyr 537
QY 662 GTAGTCAGTACGTAATCTGGAGAGGTGACAAAGCTGACCAAGGCTGACTGACCTGTACAT 721
Db 538 ValValSerTyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHis 557
```



```
Pred. No.: 5.84e-172 Length: 613
Score: 1540.00 Matches: 290
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 52.1% Indels: 0
DB: 2 Gaps: 0

US-10-825-632-6 (1-1669) x US-09-976-674-21 (1-613)

QY 2 ACAGGTACAGCAATCTCTAAAGTCACATTTTAAGATCTCAGAAATATGATGATGCTGAA 61
Db ThrGlyThrAlaAsnProLysValThrPheLysMetSerGluLeuMetIleAspAlaGlu 337
QY 62 GGAAGGATCATAGATGTCATAGATAAGAACTAATTCACCTTTTGAGATCTTATTGAA 121
Db GlyArgIleAspValIleAspLysGluLeuIleGlnProPheGluLeuPheGlu 357
QY 122 CGAGTTGATATATTCAGAGCTGGATGCTCTAGGAGAAATATGCTGGTCCATC 181
Db GlyValGluTyrIleAlaAaGAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIle 377
QY 182 CTACTAGATCGCTCCAGATCGCTACAGATAGTGTGATCTCACCTGGAATTTATTC 241
Db LeuLeuAspArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIle 397
QY 242 CCAGTAGAAGATGATGTTATGGAAGGCAGAGACTCATTGAGTCAAGTGCCTGATCTGTG 301
Db ProValGluAspAspValMetGluAaGAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerVal 417
QY 302 AGCCACTAATTTCTATGAGAAACAAACAGACATCTGGATTAATATCCATGACATCTTT 361
Db ThrProLeuIleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePhe 437
QY 362 CATGTTTTTCCCAAGTCACGAGAGAAATTTGAGTTATTTTGGCTCTGGAATGCAA 421
Db HisValPheProGlnSerHisGluGluGluLeuIleGluPheIlePheAlaSerGluCysLys 457
QY 422 ACAGTTTCCGCTCATTTATACAAATTTACATCTCTATTATTAAGGAAGCAAAATATAAACA 481
Db ThrGlyPheArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArg 477
QY 482 TCACGTGTGGGCTCGCTCCAGTGAATTTCAAGTGTCTTATCAAGAGAGATAGCA 541
Db SerSerGlyGlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAla 497
QY 542 ATTACCACTGTGAATCGGAAGTTCTTGGCCGTCAGTATTAATATCAAGTTGATGAA 601
Db IleThrSerGlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGlu 517
QY 602 GTCAAGGCTGGTATATTTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCCTGTAC 661
Db ValArgArgLeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeuTyr 537
QY 662 GTAGTCAGTTACGTAATCTCGAGAGGTGACAAAGCTCAAGCTCAAGCTGCTACAT 721
Db ValValSerTyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHis 557
QY 722 TCTTGTCTCAGTCAGCTGCTGACTCTTTATAGTAAGTATAGTAAACCAAGAGAT 781
Db SerCysCysIleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsn 577
QY 782 CCACACTGTGTCTTCAAGCTATCAAGTCTTCAAGTCTTCAAGTCTTCAAGTCTTCAAG 841
Db ProHisCysValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThr 597
QY 842 AAGGAATTTGGCCACCATTTTGGATTCA 871
Db LysGluPheTrpAlaThrIleLeuAspSer 607
```

RESULT 8

US-09-976-674-3

; Sequence 3, Application US/09976674

; Patent No. 6844180

```
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: Patent version 3.1
; SEQ ID NO 3
; LENGTH: 863
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-3
```

Alignment Scores:

Pred. No.:	1.72e-166	Length:	863
Score:	1494.50	Matches:	281
Percent Similarity:	62.9%	Conservative:	69
Best Local Similarity:	50.5%	Mismatches:	105
Query Match:	50.5%	Indels:	101
DB:	2	Gaps:	2

US-10-825-632-6 (1-1669) x US-09-976-674-3 (1-863)

```
QY 2 ACAGGTACAGCAATCTCTAAAGTCACATTTTAAGATCTCAGAAATATGATGATGCTGAA 61
Db ThrGlySerLysAsnProLysIleAlaLeuLysLeuAlaGluPheGlnThrAspSerGln 327
QY 62 GGAAGGATCATAGATGTCATAGATAAGAACTAATTCACCTTTTGAGATCTTATTGAA 121
Db GlyLysIleValSerThrGlnGluLysGluLeuValGlnProPheSerSerLeuPhePro 347
QY 122 GGAGTTGAATATATTTCCAGAGCTGGATGCTCTCAGGAGAAATATGCTTTGGTCCATC 181
Db LysValGluTyrIleAlaAaGAlaGlyTrpThrArgAspGlyLysTyrAlaTrpAlaMet 367
QY 182 CTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGGAATTTATTC 241
Db PheLeuAspArgProGlnGlnTrpLeuGlnLeuValLeuLeuProAlaLeuPheIle 387
QY 242 CCAGTAGAAGATGATGTTATGGAAGGCAGAGACTCATTGAGTCAAGTGCCTGATCTGTG 301
Db ProSerThrGluAsnGluGluGlnArgLeuAlaSerAlaArgAlaValProArgAsnVal 407
QY 302 AGCCACTAATTTATCTATGAGAAACAAACAGACATCTGGATTAATATCCATGACATCTTT 361
Db GlnProTyrValValTyrGluGluValThrAsnValTyrIleAsnValHisAspIlePhe 427
QY 362 CATGTTTTTCCCAAGTCACTC---GAAGAGAAATTTGAGTTATTTTGGCTCTCAATGC 418
Db TyrProPheProGlnSerGluGlyGluAspGluLeuLeuLeuValHisAsnGluCys 447
QY 419 AAAACAGTTTCCGTCATTTATACAAATTTACATCTATTATTAAGGAAGAAATATAA 478
Db LysThrGlyPheCysHisLeuTyrLysValThrAlaValLeuLysSerGlnGlyTyrAsp 467
QY 479 CGATCCAGTGTGGCTGCTGCTCCCAAGTATTTCAAGTGTCTTATCAAGTGTCTTATCA 538
Db TrpSerGluProPheSerProGlyGluAspGluPheLysCysProIleLysGluLeu 487
QY 539 GCAATTAACAGTGTGTAATCGGAAGTTCTTTGGCCGCGCATGATCTAATATCAAGTTGAT 598
Db AlaLeuThrSerGlyGluTrpGluValLeuAlaArgHisGlySerLysIleTrpValAsn 507
QY 599 GAACTCAGAGGCTGCTATATTTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTG 658
Db GluGluThrLysLeuValTyrPheGlnGlyThrLysAspThrProLeuGluHisLeu 527
```

```
QY 659 TACGTAGTCAGTTACCTGAGAGGTGACAAAGCTGACTGACCGTGCTACTCA 718
Db 528 TyrValValSerTyrGluAlaAlaGlyGluIleValArgLeuThrProGlyPheSer 547
QY 719 CATCTTTGTCATCAGTCAGCAGCTGTGACTTCTTTTATAAGTAAGTAACTAGTAACCAAG 778
Db 548 HisSerCysSerMetSerGlnAenPheAspMetPheValSerHisTyrSerValSer 567
QY 779 AATCCACACATGTGTCTTACAGCTATCAAGCTCTGAAGTATGAGCAACCAACTTGCAAA 838
Db 568 ThrProCysValHisValTyrLysLeuSerGlyProAspAspProLeuHisLys 587
QY 839 ACAAGGAATTTGGCCACCAATTTGGATTTCAGAGGTCTCTCTCTGACTATCTCT 898
Db 588 GlnProArgPheTrpAlaSerMetMetGluAlaLaserCysProProAspTyrValPro 607
QY 899 CCGAATTTTCCTTTTCAAGTACTACTGATTTACATTTGATGGATGCTCTCAAG 958
Db 608 ProGluIlePheHisPheHisThrArgSerAspValArgLeuTyrGlyMetIleTyrLys 627
QY 959 CCTCATGATCTACAGCTCGAAGAAATATCTACTGTCTGCTTCATATGATGGTCTCT 1018
Db 628 ProHisAlaLeuGlnProGlyLysHisProHisValLeuPheValTyrGlyGlyPro 647
QY 1019 CAG----- 1021
Db 648 GlnValGlnLeuValAenAenSerPheLysGlyIleLysTyrLeuArgLeuAenThrLeu 667
QY 1021 ----- 1021
Db 668 AlaSerLeuGlyTyrAlaValValIleAspGlyArgGlySerCysGlnArgGlyLeu 687
QY 1021 ----- 1021
Db 688 ArgPheGluGlyAlaLeuLysAenGlnMetGlyGlnValGluIleGluAspGlnValGlu 707
QY 1021 ----- 1021
Db 708 GlyLeuGlnPheValAlaGluLysTyrGlyPheIleAspLeuSerArgValAlaIleHis 727
QY 1021 ----- 1021
Db 728 GlyTrpSerTyrGlyGlyPheLeuSerLeuMetGlyLeuIleHisLysProGlnValPhe 747
QY 1022 ---GTTGCTATTCTGGGCCCCAGTCACCTCTGTGGATCTTCTATGATACAGGATACACG 1078
Db 748 LysValAlaIleAlaGlyAlaProValThrValTyrMetAlaTyrAspThrGlyTyrThr 767
QY 1079 GAAGCTTATATGGTCACCTGACAGATGAACAGGCTATTACTTGGATCTGTGGCC 1138
Db 768 GluArgTyrMetAspValProGluAenAenGlnHisGlyTyrGluAlaGlySerValAla 787
QY 1139 ATGCAAGCAGAAAGTTCCCTCTGNACCAATCGTTTACTGCTCTTACATGGTTCCCTG 1198
Db 788 LeuHisValGluLysLeuProAenGluProAenArgLeuLeuIleLeuHisGlyPheLeu 807
QY 1199 GATGAGATCTCATTGTCACATACCAGTATATTACTAGATTTTTCAGTGGCTGGA 1258
Db 808 AspGluAenValHisPheHisThrAenPheLeuValSerGlnLeuIleArgGalaGly 827
QY 1259 AAGCATATGATTTACAGATCTATCTCAGGAGACACACAGATAAGAGTTCTCTGAATCG 1318
Db 828 LysProTyrGlnLeuGlnIleTyrProAenGluArgHisSerIleArgCysProGluSer 847
QY 1319 GGAGAACATTATGAATGCTCTTTTGCACTACCTCTCAAGAAACCTT 1366
Db 848 GlyGluHisTyrGluValThrLeuLeuHisPheLeuGlnGluTyrLeu 863
```

RESULT 9

US-09-976-674-23

; Sequence 23, Application US/09976674

; Patent No. 6844180

; GENERAL INFORMATION:

```
; APPLICANT: Qi, Steve
; APPLICANT: Akineanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 23
; LENGTH: 892
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-23

Alignment Scores:
Pred. No.: 1,75e-166 Length: 892
Score: 1494.50 Matches: 281
Percent Similarity: 62.9% Conservative: 69
Best Local Similarity: 50.5% Mismatches: 105
Query Match: 50.5% Indels: 101
DB: 2 Gaps: 2

US-10-825-632-6 (1-1669) x US-09-976-674-23 (1-892)

QY 2 ACAGGTACAGCAATCTTAAAGTCACCTTTTAAGATGTCAGAAATAAATGATTGCTGAA 61
Db 337 ThrGlySerLysAenProLysIleAlaLeuLysLeuAlaGluPheGlnThrAspSerGln 356
QY 62 GGAAGATCATGATGTCATAGATAAGGAACCTAATTCACCTTTTGAGATTCTATTGAA 121
Db 357 GlyLysIleValSerThrGlnGluLysGluLeuValGlnProPheSerSerLeuPhePro 376
QY 122 GGAATTGAATATATCCAGAGCTGGATGGACTCTCTGAGGAGAAATATGCTTGTGCTCATC 181
Db 377 LysValGluTyrIleAlaArgAlaGlyTrpThrArgAspGlyLysTyrAlaTrpAlaMet 396
QY 182 CTACTAGATCGCTCCAGACTCGCCCTCAGATAGTGTGATCTCACCTGAAATTATTATC 241
Db 397 PheLeuAspArgProGlnGlnTrpLeuGlnLeuValLeuLeuProAlaLeuPheIle 416
QY 242 CCAAGTAGAGATGATGTTATGGAAAGGAGAGACTCATTGAGTCAGTCCCTGATCTGTG 301
Db 417 ProSerThrGluAenGluGluGlnArgLeuAlaSerAlaArgAlaValProArgAenVal 436
QY 302 ACGCCACTAATATTCTATCAAGAAACACAGACATCTGATAAATATCCATGACATCTTT 361
Db 437 GlnProTyrValValTyrGluGluValThrAenValTyrIleAenValHisAspIlePhe 456
QY 362 CATGTTTTCCTCCAAAGTCAC--GAAGAGGAAATTTGAGTTTATTTTTCCTCTGAAATGC 418
Db 457 TyrProPheProGlnSerGluGlyGluAspGluLeuCysPheLeuArgAlaAenGluCys 476
QY 419 AAAACAGGTTTCGTCATTTATACAAATACATCTATTTTAAAGGAAAGCAATATAAA 478
Db 477 LysThrGlyPheCysHisLeuTyrLysValThrAlaValLeuLysSerGlnGlyTyrAsp 496
QY 479 CGATCCAGTGGTGGCTGCTCTCCAAGTGAATTCAGTGAATTCAGTCTCTCAAGAGAGATA 538
Db 497 TrpSerGluProPheSerProGlyGluAspGluPheLysCysProIleLysGluIle 516
QY 539 GCAATACAGTGGTGAATGGAGGTTCTTGGCCGCGATGATCTTAATATCAAGTTGAT 598
Db 517 AlaLeuThrSerGlyGluTrpGluValLeuAlaArgHisGlySerLysIleTrpValAen 536
QY 599 GAAGTCAGAGGCTGTGATATTTTGAAGGCACAAAGACTCCCTTTAGACATCACTG 658
Db 537 GluGluThrLysLeuValTyrPheGlnGlnThrLysAspThrProLeuGluHisLeu 556
QY 659 TACGTAGTCAGTTACGTAATAATCTCGAGAGAGGTGACAAAGCTGACTGACCGTCTACTCA 718
```



```

Db      557 TyrValValSerTyrGluAlaAlaGlyGluIleValArgLeuThrThrProGlyPheSer 576
Qy      719 CATTCTTGCTGATCAGTCACAGTCGTGACTCTTTATAGTAAAGTATAGTAAACAGAG 778
Db      577 HisSerCysSerMetSerGlnAsnPheAspMetPheValSerHisTyrSerValSer 596
Qy      779 AATCACACATGTGTCTCCCTTTACAAAGTATCAAGTCTCAAGTCTCAAGTCAACCACTTGC 838
Db      597 ThrProCysValHisValTyrLysLeuSerGlyProAspAspProLeuHisLys 616
Qy      839 ACAAGGAATTTGGCCACCATTGGATTTCAGAGTCTCTCTCTGCTGCTATATCTCTCT 898
Db      617 GlnProArgPheTyrAlaSerMetMetGluAlaAlaSerCysProAspTyrValPro 636
Qy      899 CAGAAATTTCTCTTTGAAGTACTACTGGATTACATGTTGATGGGATGCTCTACAG 958
Db      637 ProGluIlePheHisPheHisThrArgSerAspValArgLeuTyrGlyMetIleTyrLys 656
Qy      959 CTTCTGATGATCAAGCTTCAAGGAAATATCTCTGCTGCTGCTGCTGCTGCTGCTGCT 1018
Db      657 ProHisAlaLeuGlnProGlyLysHisProThrValLeuPheValTyrGlyGlyPro 676
Qy      1019 CAG----- 1021
Db      677 GlnValGlnLeuValAsnAsnSerPheLysGlyIleLysTyrLeuArgLeuAsnThrLeu 696
Qy      1021 ----- 1021
Db      697 AlaSerLeuGlyTyrAlaValValAlaValIleAspGlyArgGlySerCysGlnArgGlyLeu 716
Qy      1021 ----- 1021
Db      717 ArgPheGluGlyAlaLeuLysAsnGlnMetGlyGlnValGluIleGluAspGlnValGlu 736
Qy      1021 ----- 1021
Db      737 GlyLeuGlnPheValAlaGluLysTyrGlyPheIleAspLeuSerArgValAlaIleHis 756
Qy      1021 ----- 1021
Db      757 GlyTyrSerTyrGlyGlyPheLeuSerLeuMetGlyLeuIleHisLysProGlnValPhe 776
Qy      1022 ----GTGCTATTGCTGGGCCCCAGTCATCTGTGGATCTTCTATGATACAGGATACAG 1078
Db      777 LysValAlaIleAlaGlyAlaProValThrValThrMetAlaTyrAspThrGlyTyrThr 796
Qy      1079 GAACGTTATATGGGTCAACCTGACAGAAATCAAGAGGCTTATCTTACTAGGATCTGGCC 1138
Db      797 GluArgTyrMetAspValProGluAsnAsnGlnHisGlyTyrGluAlaGlySerValAla 816
Qy      1139 ATGCAAGAGAAAGTTCCCTCTGAACCAATCGTTTACTGCTCTTACTGCTTACTGCTGCTG 1198
Db      817 LeuHisValGluLysLeuProAsnGluProAsnArgLeuLeuIleLeuHisGlyPheLeu 836
Qy      1199 GATGAGAAGTCCATTTGACATACAGTATATTTACTAGTATTTTACTAGGAGGCTGGA 1258
Db      837 AspGluAsnValHisPhePheHisThrAsnPheLeuValSerGlnLeuIleArgAlaGly 856
Qy      1259 AAGCCATATGATTACAGATCTATCTCTCAGGAGAGACACAGCATAGAGTCTCGAATCG 1318
Db      857 LysProTyrGlnLeuGlnIleTyrProAsnGluArgHisSerIleArgCysProGluSer 876
Qy      1319 GAGAACATATGAACTGATCTTTTGCATCTTTTGCATCTTTTGCATCTTTTGCATCTTTTGC 1366
Db      877 GlyGluHisTyrGluValThrLeuLeuHisPheLeuGlnGluTyrLeu 892

```

RESULT 10

US-09-976-674-27

; Sequence 27, Application US/09976674

; Patent No. 6844180

; GENERAL INFORMATION:

; APPLICANT: Qi, Steve

```

; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 27
; LENGTH: 892
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-976-674-27

```

```

Alignment Scores:      1.75e-166      Length:      892
Pred. No.:             1494.50         Matches:     281
Score:                 62.9%           Conservative: 69
Percent Similarity:    50.5%           Mismatches:  105
Best Local Similarity: 50.5%           Indels:       101
Query Match:          2                Gaps:         2
DB:

```

US-10-825-632-6 (1-1669) x US-09-976-674-27 (1-892)

```

Qy      2 ACAGGTACAGCAATCCTAAAGTCACCTTTTAAGATGTCAGAAATATGATTGATGCTGAA 61
Db      337 ThrGlySerLysAsnProLysIleAlaLeuLysLeuAlaGluPheGlnThrAspSerGln 356
Qy      62 GGAAGGATCATAGTGTCTATAGATAAGAACTAATCAACTTTTGTAGATCTATTGAA 121
Db      357 GlyLysIleValSerThrGlnGluLysGluLeuValGlnProPheSerSerLeuPhePro 376
Qy      122 GGAATTGAATATATGCCAGAGCTGGATGGACTCTCAGGGAAATATGCTGTGTCATC 181
Db      377 LysValGluTyrIleAlaArgAlaGlyThrArgAspGlyLysTyrAlaTrpAlaMet 396
Qy      182 CTACTAGATCGCTCCAGACTCGCTCAGATAGTGTGTCTCCTCAGTGAATTTATTTATC 241
Db      397 PheLeuAspArgProGlnGlnTrpLeuGlnLeuValLeuLeuProAlaLeuPheIle 416
Qy      242 CCAGTAAAGATGATGTTTATGAAAGGAGAGACTCATGTAGTCAAGTCAAGTCAAGTCAAGT 301
Db      417 ProSerThrGluAsnGluGluArgLeuAlaSerAlaArgAlaValProArgAsnVal 436
Qy      302 ACCCACTAATTTATCTATGAAGAAACAAGACATCTGGATAAATATCCATGACATCTTT 361
Db      437 GlnProTyrValValTyrGluGluValThrAsnValTrpIleAsnValHisAspIlePhe 456
Qy      362 CATGTTTTTCCCAAGTCCAC--GAAGAGAAATTTAGTTTATTTTGGCTCTGAATGC 418
Db      457 TyrProPheProGlnSerGluGlyGluAspGluLeuCysPheLeuArgAlaAsnGluCys 476
Qy      419 AAAACAGGTTTCGTCATTATATACAAATTTACATCTATTATTAAGAAAGCAATATATAA 478
Db      477 LysThrGlyPheCysHisLeuTyrLysValThrAlaValLeuLysSerGlnGlyTyrAsp 496
Qy      479 CGATCCAGTGGGGCTGCTCCTCCAAGTGAATTTCAAGTGTCTCATCAAGAGGAGATA 538
Db      497 TrpSerGluProPheSerProGlyGluAspGluPheLysCysProIleLysGluGluIle 516
Qy      539 GCAATACCAGTGGTGAATGGGAAGTCTTGGCCGGCATGGATCTATATATCAAGTTGAT 598
Db      517 AlaLeuThrSerGlyGluTrpGluValLeuAlaArgHisGlySerLysIleTrpValAsn 536
Qy      599 GAAAGTCAGAGGCTGCTGATATATTTGAAGACCAAGACTCCCTTTTAGACATCACTG 658
Db      537 GluGluThrLysLeuValTyrPheGlnGlyThrLysAspThrProLeuGluHisLeu 556
Qy      659 TAGGTAGTCAGTTACGTAATCTCTGGAGAGGTGACAGGCTGACTGACCGGTGCTACTCA 718

```

Db 557 TyrValValSerTyrGluAlaAlaGlyGluIleValArgLeuThrThrProGlyPheSer 576
QY 719 CATTTCTGTCAGTGTGACCTGCTTTTATAAGTAAGTATAGTAACAGAG 778
Db 577 HisSerCysSerMetSerGlnAsnPheAspMetPheValSerHisTyrSerValSer 596
QY 779 AATCCACACTGTGTGCTTACAGCTTATCAAGCTCTGAAGATGACCCCACTTCAGA 838
Db 597 ThrProCysValHisValTyrLysLeuSerGlyProAspAspPProLeuHisLys 616
QY 839 ACAAGAAATTTGGCCACCAATTTGGATTGAGTTCAGAGCTCTCTCTGACTATCTCT 898
Db 617 GlnProArgPheTrpAlaSerMetMetGluAlaAlaSerCysProProAspTyrValPro 636
QY 899 CCAAGAAATTTCTCTTTCAAGTACTACTGATTTACATTTGATGGATGCTCTCAAG 958
Db 637 ProGluIlePheHisPheHisThrArgSerAspValArgLeuTyrGlyMetIleTyrLys 656
QY 959 CTTCAATGATCTACAGCTGGAAGAAATATCTTACTGTCTGTTCATATATGTGTCT 1018
Db 657 ProHisAlaLeuGlnProGlyLysLysHisProThrValLeuPheValTyrGlyPro 676
QY 1019 CAG----- 1021
Db 677 GlnValGlnLeuValAsnAsnSerPheLysGlyIleLysTyrLeuArgLeuAsnThrLeu 696
QY 1021 ----- 1021
Db 697 AlaSerLeuGlyTyrAlaValValAlaValAlaAspGlyArgGlySerCysGlnArgGlyLeu 716
QY 1021 ----- 1021
Db 717 ArgPheGluGlyAlaLeuLysAsnGlnMetGlyGlnValGluIleGluAspGlnValGlu 736
QY 1021 ----- 1021
Db 737 GlyLeuGlnPheValAlaGluLysTyrGlyPheIleAspLeuSerArgValAlaIleHis 756
QY 1021 ----- 1021
Db 757 GlyTrpSerTyrGlyPheLeuSerLeuMetGlyLeuIleHisLysProGlnValPhe 776
QY 1022 ---GTTGCTATTGCTGGGGCCCACTGCTGAGTCTTCTATGATACAGGATACAG 1078
Db 777 LysValAlaIleAlaGlyAlaProValThrValTrpMetAlaTyrAspThrGlyTyrThr 796
QY 1079 GAACGTATATGGGTACCTGACCAAGTGAACAGGCTTACTTAGATCTGTGGCC 1138
Db 797 GluArgTyrMetAspValProGluAsnAsnGlnHisGlyTyrGluAlaGlySerValAla 816
QY 1139 ATGCAAGCAGAAAGTCCCTCTGACCAATCGTTCCTTACTGCTCTTACATGTTCTCTG 1198
Db 817 LeuHisValGluLysLeuProAsnGluProAsnArgLeuLeuIleLeuHisGlyPheLeu 836
QY 1199 GATGAGATGTCATTTGCACATACCATATATTTACTGAGTCTTTTACTGAGGCTGGA 1258
Db 837 AspGluAsnValHisPhePheHisThrAsnPheLeuValSerGlnLeuIleArgAlaGly 856
QY 1259 AGCCATATGATTACAGATCTATCTCAGAGAGACACAGCATAGAGTTCCTCAATCG 1318
Db 857 LysProTyrGlnLeuGlnIleTyrProAsnGluArgHisSerIleArgCysProGluSer 876
QY 1319 GGAGAAATATGAACTGTCATTTTTCACCTACTCTTCAAGAAACCTT 1366
Db 877 GlyGluHisTyrGluValThrLeuLeuHisPheLeuGlnGluTyrLeu 892

RESULT 11

US-09-976-674-33
; Sequence 33, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen

; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV

; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 33
; LENGTH: 879
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-33

Alignment Scores:

Pred. No.: 1,2e-160 Length: 879
Score: 1445.00 Matches: 274
Percent Similarity: 61.3% Conservative: 67
Best Local Similarity: 49.3% Mismatches: 101
Query Match: 48.9% Indels: 114
DB: 2 Gaps: 3

US-10-825-632-6 (1-1669) x US-09-976-674-33 (1-879)

QY 2 ACAGGTACAGCAAACTCTAAAGTCACTTTTAAGATGTCAGAAATAATGATTGATGCTGAA 61

Db 337 ThrGlySerLysAsnProLysIleAlaLeuLysLeuAlaGluPheGlnThrAspSerGln 356

QY 62 GGAAGATCATAGATGTCATAGATAAGAACTAATTCACCTTTTGAAGATTTCTATTGAA 121

Db 357 GlyLysIleValSerThrGlnGluLysGluLeuValGlnProPheSerSerLeuPhePro 376

QY 122 GGAGTTGAATATATGTCAGAGCTGGATGAGTCTCTGAGGAAATATGTTGGTCCATC 181

Db 377 LysValGluTyrIleAlaArgAlaGlyTrpThrArgAspGlyLysTyrAlaTrpAlaMet 396

QY 182 CTACTAGATCGTCCAGACTCGCTACAGATAGTGTGTGATCTCATCTGAATTTATTATC 241

Db 397 PheLeuAspArgProGlnGlnTrpLeuGlnLeuValLeuLeuProProAlaLeuPheIle 416

QY 242 CAGTAGAAGATGATGTTATGGAAGGACAGACTCAATGAGTCAGTCGCTGATTTCTGTG 301

Db 417 ProSerThrGluAsnGluGluGlnArgLeuAlaSerAlaArgAlaValProArgAsnVal 436

QY 302 AGCCACTAATTTATCTATGAAGAAACACAGACATCTGGATAAATATCCATGACATCTTT 361

Db 437 GlnProTyrValValTyrGluValThrAsnValTrpIleAsnValHisAspIlePhe 456

QY 362 CAGTTTTTCCCAAGTCAC---GAAGAGAAATTCAGTTTATTTTTCCTCTGAATGC 418

Db 457 TyrProPheProGlnSerGluGlyGluAspGluLeuCysPheLeuArgAlaAsnGluCys 476

QY 419 AAAACAGGTTTCCGTCATTTATACAAATTTACATCTATTTTTAAAGGAAAGCAAAATATAA 478

Db 477 LysThrGlyPheCysHisLeuTyrLysValThrAlaValLeuLysSerGlnGlyTyrAsp 496

QY 479 CGATCCAGTGTGGCTGCTGCTCCCAAGTGATTTCAAGTGCTCTTCAAGAGAGAGATA 538

Db 497 TrpSerGluProPheSerProGlyGluAspGluPheLysCysProIleLysGluIle 516

QY 539 GCAATTTACAGTGTGAATGGAAATTTCTTGGCGGATGATGATCTAATATCCAAGTTGAT 598

Db 517 AlaLeuThrSerGlyGluTrpGluValLeuAlaArgHisGlySer----- 531

QY 599 GAAGTCAGAGGCTGGTATATTTTGAAGGACCAAAAGACTCCCTTTTACAGCATCACCTG 658

Db 532 -----LysGlyThrLysAspThrProLeuGluHisLeu 543

QY 659 TACGTAGTACGTACGTAAATCTCTGGAGAGGTACAGAGCTGACCGTGGCTACTCA 718

Db 544 TyrValValSerTyrGluAlaAlaGlyGluIleValArgLeuThrThrProGlyPheSer 563

```
QY 719 CATTCTTCTGCATCAGTCAGCTGACTCTCTTTATAGTAAGTATAGTAACAGGAAG 778
Db 564 HisSerCysSerMetSerGlnAsnPheAspMetPheValSerHisTyrSerSerValSer 583
QY 779 AATCCACACTGTGTCTCTTTTACAACTATCAAGTCTGAAGATGACCCAACTTGCAAA 838
Db 584 ThrProCysValHisValTyrLysLeuSerGlyProAspAspProLeuHisLys 603
QY 839 ACAAGGAATTTGGGCCACCATTTTGGATTGACAGCGTCTCTTCTGACTATCTCTCT 898
Db 604 GlnProArgPheTrpAlaSerMetMetGluAlaSerCysProProAspTyrValPro 623
QY 899 CCAGAAATTTCTCTTTTGAAGTACTACTGATTTACATTTGATGGGATGCTTACAAAG 958
Db 624 ProGluIlePheHisPheHisThrArgSerAspValArgLeuTyrGlyMetIleTyrLys 643
QY 959 CCTCATGATCTACAGCCCTGGAAGAAATATCTACTGTCTGTTCATATATGTTGCTCT 1018
Db 644 ProHisAlaLeuGlnProGlyLysLysHisProThrValLeuPheValTyrGlyGlyPro 663
QY 1019 CAG----- 1021
Db 664 GlnValGlnLeuValAsnAsnSerPheLysGlyIleLysTyrLeuArgLeuAsnThrLeu 683
QY 1021 ----- 1021
Db 684 AlaSerLeuGlyTyrAlaValValIleAspGlyArgGlySerCysGlnArgGlyLeu 703
QY 1021 ----- 1021
Db 704 ArgPheGluGlyAlaLeuLysAsnGlnMetGlyGlnValGluIleGluAspGlnValGlu 723
QY 1021 ----- 1021
Db 724 GlyLeuGlnPheValAlaGluLysTyrGlyPheIleAspLeuSerArgValAlaIleHis 743
QY 1021 ----- 1021
Db 744 GlyTrpSerTyrGlyGlyPheLeuSerLeuMetGlyLeuIleHisLysProGlnValPhe 763
QY 1022 ---GTTGCTATTGCTGGGGCCCGCAGTCACCTCTGTGGATCTTCTATGATACAGGATACAG 1078
Db 764 LysValAlaIleAlaGlyAlaProValThrValTrpMetAlaTyrAspThrGlyTyrThr 783
QY 1079 GAACGTTATATGGTCACTCAGCAGATGAAACAGGCTTACTTACGATCTGTGGCC 1138
Db 784 GluArgTyrMetAspValProGluAsnAsnGlnHisGlyTyrGluAlaGlySerValAla 803
QY 1139 ATGCAAGCAGAAAGTTCCCTCTGAACCAATCGTTTACTGCTCTTACATGTTTCTCTG 1198
Db 804 LeuHisValGluLysLeuProAsnGluProAsnArgLeuLeuIleHisGlyPheLeu 823
QY 1199 GNTGAGAATGCTCCATTTTGCACATACACAGTATATTACTAGTATTTTGTAGCGGCTGGA 1258
Db 824 AspGluAsnValHisPheHisThrAsnPheLeuValSerGlnLeuIleArgAlaGly 843
QY 1259 AAGCCATATGATTACAGATCTATCTCTCAGGAGAGACACACATAGAAGTTCTCGAATCG 1318
Db 844 LysProTyrGlnLeuGlnIleTyrProAsnGluArgHisSerIleArgCysProGluSer 863
QY 1319 GAGAACATATTGAATGATCTCTTTGCACTACCTTCAAGAAACCTT 1366
Db 864 GlyGluHisTyrGluValThrLeuLeuHisPheLeuGlnGluTyrLeu 879
```

RESULT 12

US-09-976-674-35

```
; Sequence 35, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
```

```
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 35
; LENGTH: 879
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-35

Alignment Scores:
Pred. No.: 1,2e-160 Length: 879
Score: 1445.00 Matches: 274
Percent Similarity: 61.3% Conservatives: 67
Best Local Similarity: 49.3% Mismatches: 101
Query Match: 48.9% Indels: 114
DB: 2 Gaps: 3
```

US-10-825-632-6 (1-1669) x US-09-976-674-35 (1-879)

```
QY 2 ACAGGTACAGCAAAATCCTAAAGTCATCTTTTAAGATGTCAGAAATAATGATGTGCTGAA 61
Db 337 ThrGlySerLysAsnProLysIleAlaLeuLysLeuAlaGluPheGlnThrAspSerGln 356
QY 62 GGAAGATCATAGATGTCATAGATGAAGCACTAATTAATCAACCTTTTGAGATCTTATTTGAA 121
Db 357 GlyLysIleValSerThrGlnGluLysGluLeuValGlnProPheSerSerLeuPhePro 376
QY 122 GGAGTTGAATATATTGCCAGAGCTGGATGGACTCTCTGAGGGAAATATGCTTGCTCCATC 181
Db 377 LysValGluTyrIleAlaArgAlaGlyTyrThrArgAspGlyLysTyrAlaTrpAlaMet 396
QY 182 CTACTAGATCGCTCCAGACAGCTCGCTCAGATAGTGTGTGATCTCACCTGAATATTATTATC 241
Db 397 PheLeuAspArgProGlnGlnTrpLeuGlnLeuValLeuProAlaLeuPheIle 416
QY 242 CCGATGAGATGATGTTATGGAAGGAGAGAGCTCATTGAGTCAGTCCCTGATTCGTG 301
Db 417 ProSerThrGluAsnGluGluGlnArgLeuAlaSerAlaArgAlaValProArgAsnVal 436
QY 302 ACCGCACTAATATCTATCAAGAAACAAACAGACATCTCGATAAATATCCATGACATCTTT 361
Db 437 GlnProTyrValValTyrGluGluValThrAsnValTrpIleAsnValHisAspIlePhe 456
QY 362 CATGTTTTTCCCCCAAGTCAC---GAAGAGGAATTTGAGTTTATTTTTCCTCTGAATGC 418
Db 457 TyrProPheProGlnSerGluGlyGluAspGluLeuCysPheLeuArgAlaAsnGluCys 476
QY 419 AAAACAGGTTTCGTCATTTATACAAAATACATCTATTATTAAGGAAGCAAAATATAA 478
Db 477 LysThrGlyPheCysHisLeuTyrLysValThrAlaValLeuLysSerGlnGlyTyrAsp 496
QY 479 CGATCCAGTGGTGGGCTGCTCTCCAGTGAATTTCAAGTGTCTCTATCAAGAGAGAGATA 538
Db 497 TrpSerGluProPheSerProGlyGluAspGluPheLysCysProIleLysGluGluIle 516
QY 539 GCAATACCAAGTGGTGAATGGGAAGTTCTTGGCCGCGCATGGATCTTAATATCAAGTTGAT 598
Db 517 AlaLeuThrSerGlyGluTrpGluValLeuAlaArgHisGlySer----- 531
QY 599 GAAGTCAGAAGGCTGGTGTATATTTTGAAGCCACCAAGACTCCCTTTAGACATCACCTG 658
Db 532 -----LysGlyThrLysAspThrProLeuGluHisLeu 543
QY 659 TACGTAGTCAAGTACGTAATCTCTGAGAGAGGTGACAGGCTGACTGACCGTGGCTACTCA 718
Db 544 TyrValValSerTyrGluAlaAlaGlyGluIleValArgLeuThrThrProGlyPheSer 563
```


QY	1247	GTGAGGGCTGGAAGCCATATGATTTACAGATCTATCTCAGGAGACACAGCATAGA	1306
Db	261	ValArgAlaGlyLeuProTyrAspLeuGlnIleTyrProGlnGluArgHisSerIleArg	280
QY	1307	GTTCTGATCGGAGACATATGACATGCTCTTTTGCATCTTACCTTCAAGAAACCTT	1366
Db	281	ValProGluSerGlyGluHisTyrGluLeuHisSerLeuHisTyrLeuGlnGluAsnLeu	300
QY	1367	GGATCAGCTATTGCTGCTCTAAAGTGATA	1396
Db	301	GlySerArgIleAlaAlaLeuLysValIle	310
RESULT 14			
US-10-070-464-3.			
; Sequence 3, Application US/10070464			
; Patent No. 6881564			
GENERAL INFORMATION:			
; APPLICANT: ABBOTT, Catherine Anne			
; APPLICANT: GORRELL, Mark Douglas			
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES			
; FILE REFERENCE: GH-007			
; CURRENT APPLICATION NUMBER: US/10/070,464			
; CURRENT FILING DATE: 2002-03-07			
; PRIOR APPLICATION NUMBER: PCT/AU00/01085			
; PRIOR FILING DATE: 2000-09-11			
; PRIOR APPLICATION NUMBER: AU PQ5709			
; PRIOR FILING DATE: 2000-02-18			
; PRIOR APPLICATION NUMBER: AU PQ2762			
; PRIOR FILING DATE: 1999-09-10			
; NUMBER OF SEQ ID NOS: 8			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 3			
; LENGTH: 310			
; TYPE: PRT			
; ORGANISM: Homo Sapiens			
US-10-070-464-3			
Alignment Scores:			
Pred. No.:	5,39e-153	Length:	310
Score:	1377.50	Matches:	259
Percent Similarity:	83.5%	Conservative:	0
Best Local Similarity:	83.5%	Mismatches:	0
Query Match:	46.6%	Indels:	51
DB:	2	Gaps:	1
US-10-825-632-6 (1-1669) x US-10-070-464-3 (1-310)			
QY	620	TTTGAAGGCACCAAGACTCCCTTTAGACATCACCTGTACGTAGTCAGTTACGTAAT	679
Db	1	PheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSerTyrValAsn	20
QY	680	CCTGGAGAGTGCAAGGCTGACTGACCGTGGTACTCACATTTCTGCTGCATCAGTCAG	739
Db	21	ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysSileSerGln	40
QY	740	CATGTGACTCTTTTATAGTATAGTATAGTACACAGAGAAATCCACATGTGTGCTCCTT	799
Db	41	HisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeu	60
QY	800	TACAAGCTATCAAGTCTCTGAAGATGACCCCAACTTGCAAAACAAAGAAATTTTGGGCCACC	859
Db	61	TyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPheTyrAlaThr	80
QY	860	ATTTTGATTCAGCAGGCTCTCTTCTGACTATATCTCTCCAGAAATTTCTCTTTTGA	919
Db	81	IleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIlePheSerPheGlu	100
QY	920	AGTACTACTGATTTACATGTTATGGATGCTCTCAAGCTCATGATCTACAGCTGGA	979
Db	101	SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly	120
QY	980	AAGAAATATCTACTGTGCTGTTTCATATATGTTGTTCTCTCAG	1021
US-09-976-674-29			
; Sequence 29, Application US/09976674			
; Patent No. 6844180			
GENERAL INFORMATION:			
; APPLICANT: Qi, Steve			
; APPLICANT: Akinsanya, Karen			
; APPLICANT: Riviere, Pierre			
; APPLICANT: Junien, Jean-Louis			
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV			
; FILE REFERENCE: 70669			
; CURRENT APPLICATION NUMBER: US/09/976,674			
; CURRENT FILING DATE: 2001-10-12			
; PRIOR APPLICATION NUMBER: US 60/240,117			
; PRIOR FILING DATE: 2000-10-12			
; NUMBER OF SEQ ID NOS: 61			
; SOFTWARE: PatentIn version 3.1			
; SEQ ID NO 29			
; LENGTH: 832			
; TYPE: PRT			
; ORGANISM: Homo sapiens			
US-09-976-674-29			
Alignment Scores:			
Pred. No.:	2,53e-119	Length:	832
Score:	1094.50	Matches:	197
Percent Similarity:	74.1%	Conservative:	58
Best Local Similarity:	57.3%	Mismatches:	88
Query Match:	37.0%	Indels:	1
DB:	2	Gaps:	1
US-10-825-632-6 (1-1669) x US-09-976-674-29 (1-832)			
QY	2	ACAGGTACGCAATCTCTAAAGTCACCTTTTAAGATGTCAGAAATATGATGATGCTGAA	61

Db 337 ThrGlySerLysAsnProLysIleAlaLeuLysLeuAlaGluPheGlnThrAspSerGln 356
QY 62 GGAAGGATCATAGATGTATAGATAAGGAACATAATCAACCTTTTGAGATTCTATTGGAA 121
Db 357 GlyLysIleValSerThrGlnGluLysGluLeuValGlnProPheSerSerLeuPhePro 376
QY 122 GGAGTTGAATATATATCCAGAGCTGGATGGACTCTCTGAGGAGAAATATGCTTGGTCCATC 181
Db 377 LysValGluTyrIleAlaArgAlaGlyTrpThrArgAspGlyLysTyrAlaTrpAlaMet 396
QY 182 CTACTAGATCGCTCCAGACTCGCTACAGATAGTTGATCTCCTCCTGAATATTATC 241
Db 397 PheLeuAspArgProGlnGlnTrpLeuGlnLeuValLeuProAlaLeuPheIle 416
QY 242 CCAGTAGAGATGATGTTATGGAAGGAGAGACTCATTTGAGTCAGTGCCTGATTCTGTG 301
Db 417 ProSerThrGluAsnGluGlnArgLeuAlaSerAlaArgAlaValProArgAsnVal 436
QY 302 ACCCCACTAATTATCTATGAGAACAACAGACATCTGGATAATATCCATGACATCTTT 361
Db 437 GlnProTyrValValTyrGluGluValThrAsnValTrpIleAsnValHisAspIlePhe 456
QY 362 CATGTTTTCCCAAGTCAC---GAAGAGAAATGAGTTTATTTTGGCTCTGAATGC 418
Db 457 TyrProPheProGlnSerGluGlyGluAspGluLeuCysPheLeuArgAlaAsnGluCys 476
QY 419 AAACAGGTTCCGTCATTATACAAATATACATCTATTATAAGGAAAGCAAAATATAA 478
Db 477 LysThrGlyPheCysHisLeuTyrLysValThrAlaValLeuLysSerGlnGlyTyrAsp 496
QY 479 CGATCCAGTGGTGGCTCGCTCCCAAGTCATTTCAAGTGCTCTATCAAGAGGAGATA 538
Db 497 TrpSerGluProPheSerProGlyGluAspGluPheLysCysProIleLysGluLeu 516
QY 539 GCATATACAGTGGTGAATGGAAAGTCTTGGCGGCATGATCTAATATCAAGTTGAT 598
Db 517 AlaLeuThrSerGlyGluTrpGluValLeuAlaArgHisGlySerLysIleTrpValAsn 536
QY 599 GAAGTCAGAGCTGGTATATTTGAAGGACCAAGACTCCCTTTAGAGCATCACCTG 658
Db 537 GluGluThrLysLeuValTyrPheGlnGlyThrLysAspThrProLeuGluHisLeu 556
QY 659 TAGGTAGTCAGTACGTAAATCTCGAGAGGTACAGGCTGACTGACCGTGGTACTCA 718
Db 557 TyrValValSerTyrGluAlaAlaGlyGluIleValArgLeuThrThrProGlyPheSer 576
QY 719 CATTCCTGCTCAGTCAGCTGACTGCTCTTTTAAAGTAAGTATAGTAAACCAAG 778
Db 577 HisSerCysSerMetSerGlnAsnPheAspMetPheValSerHisTyrSerSerValSer 596
QY 779 AATCCACACTGTGTCCCTTTCAAGCTATCAAGTCTCTGAAGATGACCCAACTTGCAA 838
Db 597 ThrProProCysValHisValTyrLysLeuSerGlyProAspAspProLeuHisLys 616
QY 839 ACAAGGAATTTTGGGCCACCATTTTGGATTTCAGCAGGTCTCTTCTGACTATACCT 898
Db 617 GlnProArgPheTrpAlaSerMetMetGluAlaAlaSerCysProProAspTyrValPro 636
QY 899 CCAGAAATTTTCTTTTGAAGTACTACTGGATTATACATTGTATGGATGCTCTACAAG 958
Db 637 ProGluIlePheHisPheHisThrArgSerAspValArgLeuTyrGlyMetIleTyrLys 656
QY 959 CCTCATGATCTACAGCTCGAAGAAATATCTACTGTGCTTCCATATATGGTGGTCCCT 1018
Db 657 ProHisAlaLeuGlnProGlyLysLysHisProThrValLeuPheValTyrGlyPro 676
QY 1019 CAGGTTGCTATT 1030
Db 677 GlnValGlnLeu 680

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: May 2, 2006, 02:43:09 ; Search time 47.8004 Seconds
(without alignments)
2917.786 Million cell updates/sec

Title: US-10-825-632-6
Perfect score: 2957
Sequence: 1 aacaggtacagcaatccta.....aaaaaaaaaaaaaaaaaaaa 1669

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 1867569 seqs, 417829326 residues
Total number of hits satisfying chosen parameters: 3735138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -MODEL=frame+ n2p.model -DEV=xlp
-Q=/abs/ABSSWEB spool/US10825632/runat_01052006_105957_3449/app_query.fasta_1
-DB=Published Applications AA Main -QFMT=fastan -SUFFIX=n2p.rapbm
-MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1
-MATRIX=blosum62 -TRANS=human40 cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct
-THR MAX=100 -THR MIN=0 -ALIGN=15 -NODE=LOCAL -OUTFMT=ptc -NORM=ext
-HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abss02p
-USER=US10825632 @CGN 1.1 503 @runat_01052006_105957_3449 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=10 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications AA Main:
1: /cgn2_6/prodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/prodata/1/pubpaa/US08_PUBCOMB.pep:*
3: /cgn2_6/prodata/1/pubpaa/US09_PUBCOMB.pep:*
4: /cgn2_6/prodata/1/pubpaa/US10A_PUBCOMB.pep:*
5: /cgn2_6/prodata/1/pubpaa/US10B_PUBCOMB.pep:*
6: /cgn2_6/prodata/1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query No.	Score	Match	Length	ID	Description
1	2482	83.9	465	4	US-10-825-632-5	Sequence 5, Appli
2	2422	81.9	882	3	US-09-976-674-1	Sequence 1, Appli
3	2422	81.9	882	4	US-10-054-776-2	Sequence 2, Appli
4	2422	81.9	882	4	US-10-170-789-38	Sequence 38, Appli
5	2422	81.9	882	4	US-10-311-035-9	Sequence 9, Appli
6	2422	81.9	882	4	US-10-072-012-622	Sequence 622, App
7	2422	81.9	882	4	US-10-415-122-6	Sequence 6, Appli
8	2422	81.9	882	4	US-10-825-632-1	Sequence 1, Appli
9	2422	81.9	882	5	US-10-982-512-1	Sequence 1, Appli
10	2326	78.7	883	4	US-10-072-012-621	Sequence 621, App
11	2244.5	75.9	580	4	US-10-275-505-2	Sequence 2, Appli

12	2244.5	75.9	580	6	US-11-140-224-2	Sequence 2, Appli
13	1957.5	66.2	690	3	US-09-976-674-7	Sequence 7, Appli
14	1957.5	66.2	690	3	US-10-982-512-7	Sequence 7, Appli
15	1808	61.1	658	3	US-09-976-674-19	Sequence 19, Appli
16	1808	61.1	658	5	US-10-982-512-19	Sequence 19, Appli
17	1808	61.1	661	3	US-09-976-674-11	Sequence 11, Appli
18	1808	61.1	661	5	US-10-982-512-11	Sequence 11, Appli
19	1540	52.1	613	3	US-09-976-674-21	Sequence 21, Appli
20	1540	52.1	613	5	US-10-982-512-21	Sequence 21, Appli
21	1494.5	50.5	689	4	US-10-072-012-620	Sequence 620, App
22	1494.5	50.5	830	4	US-10-415-122-7	Sequence 7, Appli
23	1494.5	50.5	863	3	US-09-976-674-3	Sequence 3, Appli
24	1494.5	50.5	863	4	US-10-072-012-619	Sequence 619, App
25	1494.5	50.5	863	5	US-10-982-512-3	Sequence 3, Appli
26	1494.5	50.5	892	3	US-09-976-674-23	Sequence 23, Appli
27	1494.5	50.5	892	3	US-09-976-674-27	Sequence 27, Appli
28	1494.5	50.5	892	5	US-10-982-512-23	Sequence 23, Appli
29	1494.5	50.5	892	5	US-10-982-512-27	Sequence 27, Appli
30	1494.5	50.5	892	5	US-10-433-757-12	Sequence 12, Appli
31	1494.5	50.5	969	4	US-10-415-122-2	Sequence 2, Appli
32	1488.5	50.3	869	4	US-10-415-122-4	Sequence 4, Appli
33	1486.5	50.3	863	4	US-10-072-012-224	Sequence 224, App
34	1486.5	50.3	863	4	US-10-072-012-226	Sequence 226, App
35	1445	48.9	879	3	US-09-976-674-33	Sequence 33, Appli
36	1445	48.9	879	3	US-09-976-674-35	Sequence 35, Appli
37	1445	48.9	879	5	US-10-982-512-33	Sequence 33, Appli
38	1445	48.9	879	5	US-10-982-512-35	Sequence 35, Appli
39	1377.5	46.6	310	3	US-09-993-959-4	Sequence 4, Appli
40	1377.5	46.6	310	4	US-10-825-632-3	Sequence 3, Appli
41	1298.5	43.9	497	3	US-09-833-245-1594	Sequence 1594, Ap
42	1294.5	43.8	497	3	US-09-833-245-1593	Sequence 1593, Ap
43	1187.5	40.2	508	4	US-10-237-271-3	Sequence 3, Appli
44	1187.5	40.2	508	4	US-10-072-012-623	Sequence 623, App
45	1094.5	37.0	832	3	US-09-976-674-29	Sequence 29, Appli

ALIGNMENTS

RESULT 1
US-10-825-632-5
; Sequence 5, Application US/10825632
; Publication No. US20040191826A1
; GENERAL INFORMATION:
; APPLICANT: ABBOTT, Catherine Anne
; APPLICANT: CORRELL, Mark Douglas
; TITLE OF INVENTION: Dipeptidyl Peptidases
; FILE REFERENCE: FCSB-100-Div. 1
; CURRENT APPLICATION NUMBER: US/10/825,632
; CURRENT FILING DATE: 2004-04-15
; PRIOR APPLICATION NUMBER: US 10/070,464
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: PCT/AU00/01085
; PRIOR FILING DATE: 2000-09-11
; PRIOR APPLICATION NUMBER: AU PQ5709
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: AU PQ2762
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 465
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-825-632-5

Alignment Scores:
Pred. No.: 1.01e-245
Score: 2482.00
Percent Similarity: 100.0%
Best Local Similarity: 100.0%
Query Match: 83.9%
DB: 4
Length: 465
Matches: 465
Conservative: 0
Mismatch: 0
Indels: 0
Gaps: 0

US-10-825-632-6 (1-1669) x US-10-825-632-5 (1-465)

QY 2 ACAGGTACAGCAATCTTAAAGTCACTTTTAAAGTGTGAGAAATAATGATTGATGCTGAA 61
Db 1 ThrGlyThrAlaAsnProLysValThrPheLysMetSerGluMetIleAspAlaGlu 20
QY 62 GGAAGGATCATAGTGTATAGATAGAGAACTAAATCAACCTTTTGAGATTCATTTGAA 121
Db 21 GlyArgIleIleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGlu 40
QY 122 GGAGTTGAATATATGCCAGAGCTGGATGGACTCTCTGAGGAGAAATATGCTTGGTCCATC 181
Db 41 GlyValGluTyrIleAlaArgAlaGlyThrProGluGlyLysTyrAlaTrpSerIle 60
QY 182 CTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACTGAAATATTATTC 241
Db 61 LeuLeuAspArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIle 80
QY 242 CCGTAGAGATGATGTTATGGAAGGAGAGACTCATGTAGTCAGTCCGCTGATCTGTG 301
Db 81 ProValGluAspValMetGluArgGlnArgLeuIleGluSerValProAspSerVal 100
QY 302 ACCGCACTAATTATCATGAAGAAACAACAGACATCTGGATAAATATCCATGACATCTTT 361
Db 101 ThrProLeuIleIleTyrGluGluThrThrAspIleTyrPheIleAsnIleHisAspIlePhe 120
QY 362 CATGTTTTCCCAAGTACAGAGAGAAATTCAGTTTATTTTTGCTCTGAAATCAAA 421
Db 121 HisValPheProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLys 140
QY 422 ACAGGTTTCCTCATTTATACAAAATATACATCTATTTTAAAGAAAGCAAAATATAACGA 481
Db 141 ThrGlyPheArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArg 160
QY 482 TCCAGTGTGGCTGCTGCTCCAGTCAATGATTTCAAGTGTCTATCAAGAGGATAGCA 541
Db 161 SerSerGlyGlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAla 180
QY 542 ATTACAGGNGGAAATGGAGTCTTGGCCGATCGATCAATCAATCAAGTTGATGAA 601
Db 181 IleThrSerGlyGluTyrGluValLeuGlyArgHisGlySerAsnIleGlnValAspGlu 200
QY 602 GTCAGAGGCTGATATTTGAAGGCCACCAAGACTCCCTTTAGAGCATCACTGTAC 661
Db 201 ValArgLeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeuTyr 220
QY 662 GTAGTCAGTTAGTAAATCCTGGAGAGTGACAGGCTGACTGACCGTGGCTACTCAT 721
Db 221 ValValSerTyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHis 240
QY 722 TCTTCTGTCATCAGTCAGCATCTGACTCTCTTTATAGTAAGTATATGTAACCAAGAAAT 781
Db 241 SerCysCysIleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsn 260
QY 782 CCACACTGTGTGCTCTTTTACAAGCTATCAAGTCTCTGAGATGACCCAACTTGCAGAAACA 841
Db 261 ProHisCysValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThr 280
QY 842 AAGGAATTTTGGGCCACCATTTTGGATTTCAGCAGGCTCTTCTCTGACTATATCTCTCCA 901
Db 281 LysGluPheThrPalaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProPro 300
QY 902 GAAATTTTCTCTTTTGAAGTACTCTGGATTTACATTTGATGGGATGCTCTACAGCT 961
Db 301 GluIlePheSerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysPro 320
QY 962 CATGATCTACAGCTGGAGAAATATCTTACTGTCTGTTCATATATGGTGGTCTCAG 1021
Db 321 HisAspLeuGlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGln 340
QY 1022 GTTGCTATTGCTGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGATACAGGAA 1081
Db 341 ValAlaIleAlaGlyAlaProValThrLeuTyrPheIleAspThrGlyTyrThrGlu 360

QY 1082 CGTTTATGCGGTCACTCTGACCGAATGAACAGGGCTATTACTTAGGATCTGTGGCCATG 1141
Db 361 ArgTyrMetGlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMet 380
QY 1142 CAACGAGAAAGTTCCTCTCTGAAACCAATCGTTTACTGCTCTTACATGGTTCTCTGGAT 1201
Db 381 GlnAlaGluLysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAsp 400
QY 1202 GAGAAATGTCATTTTGCACATACCATGATATATTACTAGTGTGAGTGTGAGGCTGGAAG 1261
Db 401 GluAsnValHisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLys 420
QY 1262 CCATATGATTTACAGATCTATCTCTCAGGAGACACAGCATAAAGAGTTCCTGAAATCGGGA 1321
Db 421 ProTyrAspLeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGly 440
QY 1322 GAACATTATGAATGATCTTTTGCACTACCTTCAAGAAAACCTTGGATCACGATTGCT 1381
Db 441 GluHisTyrGluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAla 460
QY 1382 GCTCTAAAGTGATA 1396
Db 461 AlaLeuLysValIle 465

RESULT 2

US-09-976-674-1
; Sequence 1, Application US/09976674
; Patent No. US20020115843A1
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Jean-Louis
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 882
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-1

Alignment Scores:

Pred. No.:	1,97e-239	Length:	882
Score:	2422.00	Matches:	465
Percent Similarity:	82.3%	Conservative:	0
Best Local Similarity:	82.3%	Mismatches:	0
Query Match:	81.9%	Indels:	100
DB:	3	Gaps:	1

US-10-825-632-6 (1-1669) x US-09-976-674-1 (1-882)

QY 2 ACAGGTACAGCAATCTTAAAGTCACTTTTAAAGTGTGAGAAATAATGATTGATGCTGAA 61
Db 318 ThrGlyThrAlaAsnProLysValThrPheLysMetSerGluMetIleAspAlaGlu 337
QY 62 GGAAGGATCATAGTGTATAGATAGAGAACTAAATCAACCTTTTGAGATTCATTTGAA 121
Db 338 GlyArgIleIleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGlu 357
QY 122 GGAGTTGAATATATGCCAGAGCTGGATGGACTCTCTGAGGAGAAATATGCTTGGTCCATC 181
Db 358 GlyValGluTyrIleAlaArgAlaGlyTyrThrProGluGlyLysTyrAlaTrpSerIle 377
QY 182 CTACTAGATCGCTCCAGACTCGCTTACAGATAGTGTGATCTCACTGAAATATTATTC 241
Db 378 LeuLeuAspArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIle 397

QY	242	CGAGTAGAAGATCATGTTATGAAAGGACAGAGACTCATTGAGTCAGTGCCTGATTCTGTG	301	QY	1022	GTTCTATTGCTGGGCCCCAGTCACTCTGTGGATCTTTCTATGATACAGGATACACGGAA	1081
Db	398	ProValGluAspValMetGluArgGluArgGluLeuLeuGluSerValProAspSerVal	417	Db	758	ValAlaLeuAlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGlu	777
QY	302	ACGCCACTAATTATCTATGAAGAAACAAACAGACATCTCGATAAATATCCATGACATCTTT	361	QY	1082	CGTTATATGGTCACCTCCAGCAATGAACAGGGCTATTACTTAGGATCTGTGGCCATG	1141
Db	418	ThrProLeuIleIleTyrGluThrThrAspIleTrpIleAsnIleHisAspIlePhe	437	Db	778	ArgTyrMetGlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMet	797
QY	362	CATGTTTTTCCCAAGTCAAGAGGAAATGTAGTTTATTTTGTCTCTCGAATGCAAA	421	QY	1142	CAGACGAGAAAGTTCCTCCCTCGAACCAATCGTTTACTGCTCTTACATGGTTTCTCGAT	1201
Db	438	HisValPheProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLys	457	Db	798	GlnAlaGluLysPheProSerGluProAsnArgLeuLeuLeuLeuHisGlyPheLeuAsp	817
QY	422	ACAGTTTCCTCATTTATACAAATACATCTATTTTAAAGGAAACCAATATAACGA	481	QY	1202	GAGAATGTCATTTCACATACACAGTATATTACTAGTCTTTTGTAGTGGAGGCTGGAAG	1261
Db	458	ThrGlyPheArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArg	477	Db	818	GluAsnValHisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLys	837
QY	482	TCAGTGGTGGGCTGCTCTCCAGTGATTTCAAGTGTCTCATCAAGAGGAGATAGCA	541	QY	1262	CCATATGATTACAGATCTATCTCTCAGGAGACACACAGATAAGAGTTCTCGAATCGGA	1321
Db	478	SerSerGlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAla	497	Db	838	ProTyrAspLeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGly	857
QY	542	ATTACCAAGTGGTCAATGGGAAGTCTTGGCCGCATGATCTAATATCCAAAGTTGATGAA	601	QY	1322	GAACATTATGAACCTGCATCTTTTGCACTACCTTCAAGAAAACTTTGGATCAGTATTGCT	1381
Db	498	IleThrSerGlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGlu	517	Db	858	GluHisTyrGluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAla	877
QY	602	GTCAAGGCTGTATATTTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTAC	661	QY	1382	GCTCTAAAAGTGATA	1396
Db	518	ValArgArgLeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeuTyr	537	Db	878	AlaLeuLysValIle	882
QY	662	GTAGTCAGTTACGTAAATCCTCGAGAGGTGACAAAGCTGACTGACCGTGGCTACTCACAT	721	RESULT 3			
Db	538	ValValSerTyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHis	557	US-10-054-776-2			
QY	722	TCTTGTGCATCAGTCAGCACTGTGACTCTTTTATAGTAAGTATAGTAACCGAGAGAT	781	; Sequence 2, Application US/10054776			
Db	558	SerCysCysIleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsn	577	; Publication No. US20030165818A1			
QY	782	CCACACTGTGTGCTTTCACAGCTATCAAGTCTGAGATGACCCCACTGCGAACAACA	841	; GENERAL INFORMATION:			
Db	578	ProHisCysValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThr	597	; APPLICANT: Mark Robert Edbrooke			
QY	842	AAGGAAATTTGGCCACCAATTTGGATTTCAGAGTCTCTCTCTGACTATACTCTCTCCA	901	; APPLICANT: Alan Peter Lewis			
Db	598	LysGluPheTrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProPro	617	; TITLE OF INVENTION: NOVEL PROTEIN			
QY	902	GAATTTCTCTTTTGAAGTACTACTGTGATTACATTTGATGGGATGCTCTCAAGCCT	961	; FILE REFERENCE: OG1042US			
Db	618	GluIlePheSerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysPro	937	; CURRENT APPLICATION NUMBER: US/10/054,776			
QY	962	CATGATCTACAGCCTCGAAGAAATATCTACTGTGTGTTTCATATATGTGTGCTCTCAG	1021	; CURRENT FILING DATE: 2002-01-23			
Db	638	HisAspLeuGlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGln	657	; NUMBER OF SEQ ID NOS: 2			
QY	1021	-----	1021	; SOFTWARE: Patent in version 3.0			
Db	658	ValGlnLeuValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAla	677	; SEQ ID NO 2			
QY	1021	-----	1021	; TYPE: PRT			
Db	678	SerLeuGlyTyrValValValIleAspAsnArgLysSerCysHisArgLysLeuLys	697	; ORGANISM: Homo sapiens			
QY	1021	-----	1021	US-10-054-776-2			
Db	698	PheGluGlyAlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGly	717	Alignment Scores:			
QY	1021	-----	1021	Pred. No.: 1,97e-239 Length: 882			
Db	718	LeuGlnTyrLeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGly	737	Score: 2422.00 Matches: 465			
QY	1021	-----	1021	Percent Similarity: 82.3% Conservative: 0			
Db	738	TrpSerTyrGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArg	757	Best Local Similarity: 82.3% Mismatches: 0			
QY	1021	-----	1021	Query Match: 81.9% Indels: 100			
Db	738	TrpSerTyrGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArg	757	DB: 4			

242 CCAGTGAAGATGATGTTATGTAAGAAAGGCAGAGACTCATTCAGTCAGTGGCTGATCTGTG 301
Db ProValGluAspValMetGluArgGlnArgLeuIleGluSerValProAspSerVal 417
302 AGCCCACTAATTATCTATGAAGAAACAACAGACATCTGGATAAATATCCATGACATCTTT 361
Db ThrProLeuIleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePhe 437
362 CATGTTTTTCCCCAAAGTCACGAAGAGAAATTCAGTTTATTTTTCCTCTGAAATGCAA 421
Db HisValPheProGlnSerHisGluGluIleGluPheIlePheAlaSerGluCysIys 457
422 ACAGGTTTCCGTCATTATACAAATATACATCTATTTTAAAGAAAGCAAAATATAACGA 481
Db ThrGlyPheArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysIysArg 477
482 TCCAGTGGTGGCTGCCCTCCAGTGATTTCAAGTGTCCTATCAAGAGGAGATAGCA 541
Db SerSerGlyGlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAla 497
542 ATTACAGTGGTGAATGGAGTCTTGGCCGCATCGATCTAATATCCAGTTGATGAA 601
Db IleThrSerGlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGlu 517
602 GTCAGAAGCTGGTATTTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTAC 661
Db ValArgArgLeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeuTyr 537
662 GTAGTCAGTTACGTAAATCCTGGAGAGGTGACAAAGGCTGACTGACCGTGGCTACTCACAT 721
Db ValValSerTyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHis 557
722 TCTGCTGCATCAGTCAGCTGTGACTTCTTTTAAAGTAAATAGTATAGTAACACAGAGAAAT 781
Db SerCysCysIleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsn 577
782 CCACACTGTGTCTCCTTTACAAGCTATCAAGTCTCTGAAGTACCCCACTTGCAGAAACA 841
Db ProHisCysValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThr 597
842 AAGGAATTTGGCCACCATTTTGATTCAGCAGCTCTCTCTCTGACTATATCTCTCCA 901
Db LysGluPheTrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProPro 617
902 GAAATTTCTCTTTTGAAGTACTACTGGATTTACATTTGATGGATGCTCTACAGCT 961
Db GluIlePheSerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysPro 637
962 CATGATCTACGCTGGAAGAAATATCTACTGCTGTGTTTCATATATGTTGCTCTCAG 1021
Db HisAspGluProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGln 657
1021 ----- 1021
658 valGlnLeuValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAla 677
1021 ----- 1021
678 SerLeuGlyTyrValValValValIleAspAsnArgGlySerCysHisArgGlyLeuLys 697
1021 ----- 1021
698 PheGluGlyAlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGly 717
1021 ----- 1021
718 LeuGlnTyrLeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGly 737
1021 ----- 1021
738 TrpSerTyrGlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArg 757
1022 GTTGCTATTGCTGGGCCCCAGTCACCTCTGTGGATCTTCTATGATACAGGATACACGGA 1081

Db ValAlaIleAlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGlu 777
1082 CGTTATATGGGTACCCCTGACCAAGTAAACAGGGCTATTACTTAGGATCTGTGGCCATG 1141
Db ArgTyrMetGlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMet 797
1142 CAAGCAGAAAAAGTTCCTCTCTGAAACCAATCGTTTACTGCTCTTACATGGTTTCTGGAT 1201
Db GlnAlaGluLysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAsp 817
1202 GAGAATGTCATTTTGCATACATACAGTATATTAAGTATTTTGTAGTGGGCTGGAAG 1261
Db GluAsnValHisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLys 837
1262 CCATATGATTACAGATCTATCTCCAGAGAGACACAGCATAAGAGTTCCTGAATCGGA 1321
Db ProTyrAspLeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGly 857
1322 GAACATTATGAAGTGCATCTTTTGCACCTACCTTCAAGAAACCTTGGATCAGTATTGCT 1381
Db GluHisTyrGluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAla 877
1382 GCTCTAAAAGTCATA 1396
Db AlaLeuLysValIle 882
RESULT 4
US-10-170-789-38
; Sequence 38, Application US/10170789
; Publication No. US20030180930A1
; GENERAL INFORMATION:
; APPLICANT: Rachel E. Meyers
; APPLICANT: Olandt, Peter J.
; APPLICANT: Kapeller-libermann, Rosana
; APPLICANT: Curtis, Rory A. J.
; APPLICANT: Williamson, Mark
; APPLICANT: Weich, Nadine
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASE, PHOSPHATASE,
; FILE REFERENCE: 10448-191001
; CURRENT APPLICATION NUMBER: US/10/170,789
; PRIOR FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: US 09/797,039
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: PCT/US01/06525
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/186,061
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 09/882,166
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: PCT/US01/19269
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/212,078
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 09/934,406
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: PCT/US01/26052
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/226,740
; PRIOR FILING DATE: 2000-08-21
; PRIOR APPLICATION NUMBER: US 09/861,801
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: PCT/US01/16549
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: US 60/205,508
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: US 09/801,267
; PRIOR FILING DATE: 2001-03-06
; PRIOR APPLICATION NUMBER: PCT/US01/07138
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: US 60/187,454
; PRIOR FILING DATE: 2000-03-07

```
; PRIOR APPLICATION NUMBER: US 09/829,671
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: PCT/US01/40483
; PRIOR FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: US 60/197,508
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: US 09/961,721
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: PCT/US01/29904
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: US 60/235,023
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US 10/045,367
; PRIOR FILING DATE: 2001-11-07
; PRIOR APPLICATION NUMBER: US 60/246,561
; PRIOR FILING DATE: 2000-11-07
; PRIOR APPLICATION NUMBER: US 09/801,275
; PRIOR FILING DATE: 2001-03-06
; PRIOR APPLICATION NUMBER: PCT/US01/07074
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: US 60/187,420
; PRIOR FILING DATE: 2000-03-07
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 38
; LENGTH: 882
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-170-789-38

Alignment Scores:
Pred. No.: 1,976-239 Length: 882
Score: 2422.00 Matches: 465
Percent Similarity: 82.3% Conservative: 0
Best Local Similarity: 82.3% Mismatches: 0
Query Match: 81.9% Indels: 100
DB: 4 Gaps: 1

US-10-825-632-6 (1-1669) x US-10-170-789-38 (1-882)

QY 2 ACAGGTACAGCAATCTTAAGTCACTTTTAAGATGTCAGAAATAATGATGTCGAA 61
DB 318 ThrGlyThrAlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGlu 337
QY 62 GGAAGGATCATAGATCATAGATAGGAACCTAATCAACCTTTTGAGATTCATTGAA 121
DB 338 GlyArgIleIleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGlu 357
QY 122 GAGTTGAATATATATCCAGAGCTGGATGACTCCTGAGGGAAATATGCTTGTCATC 181
DB 358 GlyValGluTyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIle 377
QY 182 CTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAAATATTTATC 241
DB 378 LeuLeuAspArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIle 397
QY 242 CCAGTACAGATCATGTTATGGAAGGACAGACTCATGAGTCAGTGCCTGATTCGTG 301
DB 398 ProValGluAspValMetGluArgGlnArgLeuIleGluSerValProAspSerVal 417
QY 302 ACGCCATAATTATCTATGAAGAAACAACAGACATCTGGATAAATATCCATGATCTTT 361
DB 418 ThrProLeuIleIleTyrGluIleThrThrAspIleTrpIleAsnIleHisAspIlePhe 437
QY 362 CATGTTTTCCTCCCAAGTACGAAAGGAAATGATTTATTTTGGCTCTGAAATGCAAA 421
DB 438 HisValPheProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLys 457
QY 422 ACAGTTTCGTCATTTATCAAAATATACATCTATTTAAAGGAAGCAAAATATAACGA 481
DB 458 ThrGlyPheArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysArg 477
QY 482 TCCAGTGTGGCTCGCTGCTCCCAAGTGTATTTCAAGTGTCTTATCAAGAGGAGATAGCA 541
```

```
DB 478 SerSerGlyGlyLeuProAlaProSerAspPheLysCysProIleLysGluIleAla 497
QY 542 ATTACAGTGGTCAATGGGAAGTCTTGGCCGGATGATCTAATATATCAAGTTGATGAA 601
DB 498 IleThrSerGlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGlu 517
QY 602 GTCAGAAAGCTGGTATATATTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACCTGTAC 661
DB 518 ValArgArgLeuValTyrPheGluGlyThrLysAspSerProLeuGluHisIleLeuTyr 537
QY 662 GTAGTCAGTTTACGTAATCTCGAGAGGTGACAAAGCTGACTGACCGGTGCTACTCAT 721
DB 538 ValValSerTyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHis 557
QY 722 TCTGTGTCATCAGTCAGTCAGTGTGACTTCTTTTATAGTAAGTATAGTAACCAAGAAAT 781
DB 558 SerCysCysIleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsn 577
QY 782 CCACACTGTGTGCTTCCTTTTACAAGCTATCAAGTCTCTGAAGATGACCCAACTTGCAAAACA 841
DB 578 ProHisCysValSerLeuTyrLysLeuSerProGluAspAspProThrCysLysThr 597
QY 842 AAGGAATTTTGGCCACCACTTTTGGATTCAGCAGGTCTCTTCTCTGACTATCTCTCCA 901
DB 598 LysGluPheTrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProPro 617
QY 902 GAAATTTTCTCTTTGAAAGTACTACTGTGATTTACATTTGATGGATCTCTACAAGCCT 961
DB 618 GluIlePheSerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysPro 637
QY 962 CATGATCTACAGCTCGAAAGAAATATCTACTGTGCTGTTTCATATATATGTTGGTCTCAG 1021
DB 638 HisAspLeuGlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGln 657
QY 1021 ----- 1021
DB 658 ValGlnLeuValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAla 677
QY 1021 ----- 1021
DB 678 SerLeuGlyTyrValValValValIleAspAsnArgGlySerCysHisArgGlyLeuLys 697
QY 1021 ----- 1021
DB 698 PheGluGlyAlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGly 717
QY 1021 ----- 1021
DB 718 LeuGlnTyrLeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGly 737
QY 1021 ----- 1021
DB 738 TrpSerTyrGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArg 757
QY 1022 GTTCTATTCTGGGGCCCCAGTCACTCTGCGATCTCTATGATACAGGATACACGAA 1081
DB 758 ValAlaIleAlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGlu 777
QY 1082 CGTTATATGGTGCACCTGACACCAAGTGAACAGGGCTATTACTAGGATCTGTGGCCATG 1141
DB 778 ArgTyrMetGlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMet 797
QY 1142 CAACGACAAAGATTTCCCTCTGAAACCAATCGTTTACTCTCTTACATGGTTTCTCGAT 1201
DB 798 GlnAlaGluLysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAsp 817
QY 1202 GAGAAATGTCATTTTGCATACATACAGTATATTACTAGTGTGTTTGTAGTGGGCTGGAAG 1261
DB 818 GluAsnValHisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLys 837
QY 1262 CCATATGATTTACAGATCTATCTCAGGAGACACAGCATAGAGTTCTCTCAATCGGA 1321
```

```
Db      838  ProTyrAspLeuGluInileTyrProGlnGluAArgHisSerIleAArgValProGluSerGly 857
Qy      1322  GAACATTATCAACTGCATCTTTTGGCACTACCTTCAAGAAACCTTGGATCAGCTATTGCT 1381
Db      858  GluHisTyrGluLeuHisLeuLeuHisTyrLeuGlnGluAenLeuGlySerArgIleAla 877
Qy      1382  GCTCTAAAGCTGATA 1396
Db      878  AleLeuLysValIle 882

RESULT 5
US-10-311-035-9
; Sequence 9 Application US/10311035
; Publication N US20040023243A1
; GENERAL INFORMATION: INCYTE GENOMICS, INC.
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: YUE, Henry
; APPLICANT: ELIOTT, Vicki
; APPLICANT: GANDHI, Ameena R.
; APPLICANT: LAL, Preeti
; APPLICANT: AU-YOUNG, Janice
; APPLICANT: TRIBOULEY, Catherine M.
; APPLICANT: DEBEAUNE, Angelo M.
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: NGUYEN, Dannie B.
; APPLICANT: LEE, Ernestine A.
; APPLICANT: HAFALIA, April
; APPLICANT: KHAN, Farrah A.
; APPLICANT: CHAWLA, Narinder K.
; APPLICANT: YAO, Monique G.
; APPLICANT: LU, Dyung Aina M.
; APPLICANT: ARVIZU, Chandra S.
; APPLICANT: TANG, Y. Tom
; APPLICANT: WALSH, Roderick T.
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: LU, Yan
; APPLICANT: RAMKUMAR, Jayalaximi
; APPLICANT: XU, Yuming
; APPLICANT: REDDY, Roopa
; APPLICANT: DAS, Depodriya
; APPLICANT: KEARNEY, Liam
; APPLICANT: KALLICK, Deborah A.
; TITLE OF INVENTION: Proteases
; FILE REFERENCE: PI-012356
; CURRENT APPLICATION NUMBER: US/10/311,035
; CURRENT FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: 60/212,336; 60/213,955; 60/215,396; 60/216,821; 60/218,946
; PRIOR FILING DATE: 2000-06-16; 2000-06-22; 2000-06-29; 2000-07-07; 2000-07-14
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PERL Program
; SEQ ID NO 9
; LENGTH: 882
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20040023243A1 7160544CD1
US-10-311-035-9

Alignment Scores:
Pred. No.: 1,97e-239 Length: 882
Score: 2422.00 Matches: 465
Percent Similarity: 82.3% Conservative: 0
Best Local Similarity: 82.3% Mismatches: 0
Query Match: 81.9% Indels: 100
DB: 4 Gaps: 1

US-10-825-632-6 (1-1669) x US-10-311-035-9 (1-882)
Qy      2  ACAGTACAGCAATCCTTAAGTCACATTTTAAGATGTCAGAAATATGATTCAGTCAA 61
Db      318  ThrGlyThrAlaAsnProLysValThrPheLysMetSerGluIleMetIleAlaGlu 337

62  CGAAGGATCATAGATGTCATAGATTAAGAACTAAATCAACCTTTTGAGATTCATTGAA 121
338  GlyArgIleIleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGlu 357
122  GGATTTGAATATATTTGCCAGAGCTGGATCGATCTCTGAGGAAAAATATGCTTGGTCCATC 181
358  GlyValGluTyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIle 377
182  CTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGTGATCTCACTACCTGAATTTATTATC 241
378  LeuLeuAspArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIle 397
242  CCAGTAGAAGATGATGTTATGGAAGGCAGAGACTCATTTGAGTCAGTCCCTGATTCGTGTG 301
398  ProValGluAspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerVal 417
302  ACGCCACTAATTATCTATGAAGAAACAACAGACATCTGGATAAATATCCATGACATCTTT 361
418  ThrProLeuIleIleTyrGluGluThrThrAspIleTyrPheAsnIleHisAspIlePhe 437
362  CATGTTTTTCCCAAGTCAAGAGGAAATTTAGTTTATTATTTTGGCTCTGAAATGCAAA 421
438  HisValPheProGlnSerHisGluGluGluIleGluPheIleAspSerGluCysLys 457
422  ACAGGTTTCGTCATTTATACAAAATTACATCTATTTTAAAGAAAGCAATATAAACA 481
458  ThrGlyPheArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArg 477
482  TCCAGTGGTGGCTGCTGCTCCAAAGTATTCAGTGTCTCTATCAAGAGAGATAGCA 541
478  SerSerGlyLeuLeuProIleProSerAspPheLysCysProIleLysGluGluIleAla 497
542  ATTACCAAGTGGTGAATGGGAAGTTCTTGGCCGGCATGGATCTAATATCAAGTTGATGAA 601
498  IleThrSerGlyLeuTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGlu 517
602  GTCAGAAGCTGTATATTTTGAAGGCCACCAAGACTCCCTTTTAGAGCATCCTCTAC 661
518  ValArgArgLeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeuTyr 537
662  GTAGTCAGTTACGTAAATCCTGGAGAGGTGACAAAGGTGACTGACCGTGGCTACTCATCAT 721
538  ValValSerTyrValAsnProGlyValThrArgLeuThrAspArgGlyTyrSerHis 557
722  TCTTGTGTCATCAGTCAGCTGTGACTTCTTTTAAAGTAAGTATAGTAAACCAAGAAAT 781
558  SerCysCysIleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsn 577
782  CCACACTGTGTCTCCCTTTTACAAGCTATCAAGTCTCTGAAGATGACCCAACTTGCAAAACA 841
578  ProHisCysValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThr 597
842  AAGGAATTTTGGGCCACCATTTTGGATTACAGCAGTCTCTTCTCTGACTATATCTCTCCA 901
598  LysGluPheTrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProPro 617
902  GAAATTTTCTCTTTTGAAGTACTACTGATTTTACATTTGATGGATGCTCTCAAGCCT 961
618  GluIlePheSerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysPro 637
962  CATGATCTACAGCTGGAAAGAAATATCCTACTGCTGTGTCTATATATGCTGCTCTCAG 1021
638  HisAspLeuGlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGln 657
1021  ----- 1021
658  ValGlnLeuValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAla 677
1021  ----- 1021
678  SerLeuGlyTyrValValValValIleAspAsnArgGlySerCysHisArgGlyLeuLys 697
1021  ----- 1021
```

```
Db 698 PheGluGlyAlaPheLysTyrLysMetGlyGlnIleuLeuAspGlnValGluGly 717
QY 1021 ----- 1021
Db 718 LeuGlnTyrLeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGly 737
QY 1021 ----- 1021
Db 738 TrpSerTyrGlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArg 757
QY 1022 GTTGTATTGTGGGGCCCCAGTCACCTCTGTGTGATCTTCTATGATACAGGATACACGGAA 1081
Db 758 ValAlaIleAlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGlu 777
QY 1082 CTTTATATGGTCCACCTGACCAAGATGAACAGGGCTATTACTAGGATCTGTGGCCATG 1141
Db 778 ArgTyrMetGlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMet 797
QY 1142 CAAGCAGAAAGTCCCTCTGAACCAATCGTTTACTGCTCTTACATGGTTTCTGGAT 1201
Db 798 GlnAlaGluLysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAsp 817
QY 1202 GGAATGTCATTTTGCACATACCATATATTACTGAGTTTCTTACTGAGGGCTGGAAG 1261
Db 818 GluAsnValHisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLys 837
QY 1262 CCATATGATTTTACAGATCTCTCTCAGGAGAGACAGCATAGAGTTCTCGAATCGGGA 1321
Db 838 ProTyrAspLeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGly 857
QY 1322 GAACATATGAACTGCATCTTTTGCACCTACCTTCAAGAAAACCTTGGGATCAGTATTGCT 1381
Db 858 GluHisTyrGluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAla 877
QY 1382 GCTCTAAAGTGATA 1396
Db 878 AlaLeuLysValIle 882
RESULT 6
US-10-072-012-622
; Sequence 622, Application US/10072012
; Publication No. US20040033493A1
; GENERAL INFORMATION:
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Patturajan, Meera
; APPLICANT: Shinkets, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangolli, Esha
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Hastelli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Taupier Jr, Raymond J.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Colman, Steven D.
; APPLICANT: Wolenc, Adam R.
; APPLICANT: Pena, Carol E. A
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Drosse, William M.
; APPLICANT: Alsbrook II, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-258
; CURRENT APPLICATION NUMBER: US/10/072,012
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/265,102
; PRIOR FILING DATE: 2001-01-30
```

```
; PRIOR APPLICATION NUMBER: 60/265,514
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/266,406
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/266,767
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/267,057
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/266,975
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/267,459
; PRIOR FILING DATE: 2001-02-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1391
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 622
; LENGTH: 882
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-072-012-622
Alignment Scores:
Pred. No.: 1,97e-239 Length: 882
Score: 2422.00 Matches: 465
Percent Similarity: 82.3% Conservative: 0
Best Local Similarity: 82.3% Mismatches: 0
Query Match: 81.9% Indels: 100
DB: 4 Gaps: 1
```

```
US-10-825-632-6 (1-1669) x US-10-072-012-622 (1-882)
QY 2 ACAGTACAGCAAAATCCTAAAGTCACCTTTTAAGATCTCAGAAATAATGATTGATCTGAA 61
Db ThrGlyThrAlaAsnProLysValThrPheLysMetSerGluLeuMetIleAspAlaGlu 337
QY 62 GGAAGGATCATAGATGTCTAGATAGGAACCTAACTCAACCTTTTTCAGATTCTATTGAA 121
Db GlyArgIleIleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGlu 357
QY 122 GGAGTTGAATATATTGCCAGAGCTGGATGGACTCTCTGAGGAAATAATGCTTGGTCCATC 181
Db GlyValGluTyrIleAlaArgAlaGlyTyrThrProGluGlyLysTyrAlaTrpSerIle 377
QY 182 CTACTAGATCGCTCCAGACTCGCCTACAGATAGTGTTCATCTACCTGGAATTATTATC 241
Db LeuLeuAspArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIle 397
QY 242 CCAGTAGAAGATGATGTATTGGAAGCAGACAGACTCATTGACTGCTGCTGATCTCTG 301
Db ProValGluAspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerVal 417
QY 302 ACGCCACTAATTATCTATGAAGAAACAAGACAGCATCTGATAAATATCATGACATCTTT 361
Db ThrProLeuIleIleTyrGluIuThrThrAspIleTrpIleAsnIleHisAspIlePhe 437
QY 362 CATGTTTTCCTCCAAAGTCACGAAGAGAAATAGATTATTATTTTTCCTCTGAATGCAAA 421
Db HisValPheProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLys 457
QY 422 ACAGTTTCCGTCATTTTATACAAATTAATCATCTATTATTTTAAAGGAAGCAAAATATAACGA 481
Db ThrGlyPheArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArg 477
QY 482 TCCAGTGTGGCTGCTGCCATGATTTTCAAGTGATTTTCAAGTCTTATCAAGAGAGATAGCA 541
Db SerSerGlyGlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAla 497
```

QY	542	AThACAGTGGTGAATGGGAAGTCTTGGCCGGCAGATCTAATATATCCAAAGTTGATGAA	601	Db	858	GIuHiTyGIuLeuHisLeuLeuHisTyrrLeuGIuGIuAenLeuGlySerArgGileAla	877
Db	498	IlEthrSerGIuTrpGIuValLeuGIuArgHISGlySerAenIleGIuValAspGlu	517	QY	1382	GCTCTAAAGTGATA	1396
QY	602	GTCAGAGGCTGTATATTTGAAGCACCAAGACTCCCTTTAGAGCATCACCTGTAC	661	Db	878	AlaLeuLysValIle	882
Db	518	ValArgArgLeuValTyrrPheGluGIuThyrrLysAspSerProLeuGIuHISHisLeuTyrr	537	RESULT 7			
QY	662	GTAGTCAGTTACGTAATCTCGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACAT	721	US-10-415-122-6			
Db	538	ValValSerTyrrValAenProGlyGIuValThrArgLeuThrAspArgGlyTyrrSerHis	557	; Sequence 6, Application US/10415122			
QY	722	TCTGTGTCATCAGTCAGACTGTGACTTCTTTATAAGTAAGTATAGTAACCAAGAAAT	781	; Publication No. US20040053369A1			
Db	558	SerCysCysIleSerGlnHisCysAspPhePheIleSerLysTyrrSerAenGlnLysAen	577	; GENERAL INFORMATION:			
QY	782	CCACACTGTGTGCTTACAGCTATCAAGTCTCGAAGTACCAAGTTCGCAAAACA	841	; APPLICANT: THE UNIVERSITY OF SYDNEY			
Db	578	ProHisCysValSerLeuTyrrLysLeuSerSerProGluLysAspProThrCysLysThr	597	; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES			
QY	842	AAGGAATTTGGCCACCACTTTGGATTCCAGAGTCTCTTCCGTACTATCTCTCCA	901	; FILE REFERENCE: PF15217			
Db	598	LysGluPheThrAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrrThrProPro	617	; CURRENT APPLICATION NUMBER: US/10/415,122			
QY	902	GAATTTTCTTTGAAGTACTACTGATTTACATTTGATGGATGCTCTCAAGCCT	961	; CURRENT FILING DATE: 2003-08-07			
Db	618	GluIlePheSerPheGluSerThrThrGlyPheThrLeuTyrrGlyMetLeuTyrrLysPro	637	; NUMBER OF SEQ ID NOS: 8			
QY	962	CATGATCTACAGCTCGAAGAAATATCTACTGTGCTGTTTCATATATGTGTGCTCAG	1021	; SOFTWARE: PatentIn version 3.1			
Db	638	HisAspLeuGlnProGlyLysLysTyrrProThrValLeuPheIleTyrrGlyGlyProGln	657	; SEQ ID NO 6			
QY	1021	-----	1021	; LENGTH: 882			
Db	658	ValGlnLeuValAenAenArgPheLysGlyValLysTyrrPheArgLeuAenThrLeuAla	677	; TYPE: PRT			
QY	1021	-----	1021	; ORGANISM: Homo sapiens			
Db	678	SerLeuGlyTyrrValValValIleAspAenArgGlySerCysHisArgGlyLeuLys	697	US-10-415-122-6			
QY	1021	-----	1021	Alignment Scores:			
Db	698	PheGluGlyAlaPheLysTyrrLysMetGlyGlnIleGluIleAspGlnValGluGly	717	Pred. No.: 1,97e-239 Length: 882			
QY	1021	-----	1021	Score: 2422.00 Matches: 465			
Db	718	LeuGlnTyrrLeuAlaSerArgTyrrAspPheIleAspLeuAspArgValGlyIleHisGly	737	Percent Similarity: 82.3% Conservative: 0			
QY	1021	-----	1021	Best Local Similarity: 82.3% Mismatches: 0			
Db	738	TrpSerTyrrGlyTyrrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArg	757	Query Match: 81.9% Indels: 100			
QY	1022	GTTGCTATTCTGGGGCCCACTCTGTGGATCTTCTATGATACAGGATACACGGAA	1081	Gaps: 1			
Db	758	ValAlaIleAlaGlyAlaProValThrLeuThrPheTyrrAspThrGlyTyrrThrGlu	777	US-10-825-632-6 (1-1669) x US-10-415-122-6 (1-882)			
QY	1082	CGTTATATGGGTACCCCTGACCGAAGTAACAGGGCTATTACTTAGATCTGTGGCCATG	1141	QY	2	ACAGGTACAGCAATCCTAAAGTCACTTTTAAGATGTGCAGAAATAATGATGTGCTGAA	61
Db	778	ArgTyrrMetGlyHisProAspGlnAenGluGlnGlyTyrrTyrrLeuGlySerValAlaMet	797	Db	318	ThrGlyThrAlaAenProLysValThrPheLysMetSerGluIleMetIleAspAlaGlu	337
QY	1142	CAACAGAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGTTTCTCGAT	1201	QY	62	GGAGGATCATAGATCTCATAGATAAGGAACCTAAATCAACCTTTTGAGATTCTATTGAA	121
Db	798	GlnAlaGluLysPheProSerGluProAenArgLeuLeuLeuHisGlyPheLeuAsp	817	Db	338	GlyArgIleIleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGlu	357
QY	1202	GAGAAATCTCCATTTTGACATACCACTATATTACTAGTCTTTTGTAGAGGCTGGAAAG	1261	QY	122	GGAGTTGAATATATTCAGAGAGCTGGAGTCTCTCGAGGAAAATATGCTTGTGCTCATC	181
Db	818	GluAenValHisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLys	837	Db	358	GlyValGluTyrrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrrAlaTrpSerIle	377
QY	1262	CCATATGATTTACAGATCTATCTCTCAGGAGACACACAGATAAGAGTTTCTGAAATCGGGA	1321	QY	182	CTACTAGATCCCTCCAGACTCGCCTACAGATAGTGTGATCTCACCTGAAATATTATC	241
Db	838	ProTyrrAspLeuGlnIleTyrrProGlnGluArgHisSerIleArgValProGluSerGly	857	Db	378	LeuLeuAspArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIle	397
QY	1322	GAACATTATGAACCTGTCATCTTTTGCACTACCTTCAAGAAACCTTGGATCACGTTATGCT	1381	QY	242	CCAGTAGAAGATCATCTTATGAAAGGAGAGACTCATTGAGTCAGTCCCTGATTCTGTG	301

498	Db	498	IleThrSerGlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGlu	517
602	QY	602	GTCAAGAGCGTGTATATTTTGAAGGACCAAGACTCCCTTTAGACATCACCTGTAC	661
518	Db	518	ValArgArgLeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeuTyr	537
662	QY	662	GTAGTCAGTTACGTAATAATCCTGGAGAGGTGACAGGCTGACTGACCGTGGCTACTCACAT	721
538	Db	538	ValValSerTyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHis	557
722	QY	722	TCTCTCTGCATCAGTCAGCACTGTGACTCTCTTTATAAGTAAGTATAGTAACGACAGAAT	781
558	Db	558	SerCysCysIleSerGlnHisCysAspPheIleSerLysLysSerAsnGlnLysAsn	577
782	QY	782	CCACACTGTGTCTCCCTTTTACAAGCTATCAAGTCTTGAAGATGACCCAACTTGCAAAACA	841
578	Db	578	ProHisCysValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThr	597
842	QY	842	AAGAAATTTTGGCGCCACCATTTTGGATTGAGAGGTCTCTTCTCTGACTATACCTCTCCA	901
598	Db	598	LysGluPheTpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProPro	617
902	QY	902	GAATTTTCTCTTTTGAAGTACTACTGGAATTTACATTGTATGGCATCTCTACAAGCCT	961
618	Db	618	GluIlePheSerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysPro	637
962	QY	962	CATGATCTACAGCCTGGAAGAAATATCTTACTGTCTGTCTTATATATGTTGGTCTCTCAG	1021
638	Db	638	HisAspLeuGlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGln	657
1021	QY	1021	-----	1021
658	Db	658	ValGlnLeuValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAla	677
1021	QY	1021	-----	1021
678	Db	678	SerLeuGlyTyrValValValValIleAspAsnArgGlySerCysHisArgGlyLeuLys	697
1021	QY	1021	-----	1021
698	Db	698	PheGluGlyAlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGly	717
1021	QY	1021	-----	1021
718	Db	718	LeuGlnTyrLeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGly	737
1021	QY	1021	-----	1021
738	Db	738	TrpSerTyrGlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArg	757
1022	QY	1022	GTTCGTATTGCTGGGGCCCCAGTCACCTGTGGATCTCTATCATACAGGATACACGGAA	1081
758	Db	758	ValAlaIleAlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGlu	777
1082	QY	1082	CGTTATATGGGTCACCCTGCAGCAATGAACAGGGCTATTACTTAGGATCTGTGGCCATG	1141
778	Db	778	ArgTyrMetGlyHisProAspGlnAsnGlnGlyTyrTyrLeuGlySerValAlaMet	797
1142	QY	1142	CAAGCAGAAAAGTCCCTCTGAACCAAACTCGTTTACTGCTCTTACATGGTTTCTCGAT	1201
798	Db	798	GlnAlaGlyLysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAsp	817
1202	QY	1202	GAGAATGTCATTTTGCACATACAGATATATTACTGAGTTTTTTAGTCAGGCGCTGGAAG	1261
818	Db	818	GluAsnValHisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLys	837
1262	QY	1262	CCATATGATTTACAGATCTATCTCTCAGGAGAGACACAGCATAAAGATTCTCTGAATCGGGA	1321
838	Db	838	ProTyrAspLeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGly	857
1322	QY	1322	GAACATTATGAATCGACTCTTTTGCATCTCTTCAGAAAACCTTGGATCAGTATTGCT	1381
858	Db	858	GluHisTyrGluLeuHisLeuLeuHisTyrLeuGlnGlnLeuGlySerArgIleAla	877

```

Qy      1382  GCTCTAAAGTGATA 1396
      |||||
Db      878  AlaleuylsVaille 882

RESULT 8
US-10-825-632-1
; Sequence 1, Application US/10825632
; Publication No. US2004019B26A1
; GENERAL INFORMATION:
; APPLICANT: ABBOTT, Catherine Anne
; APPLICANT: CORRELL, Mark Douglas
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES
; FILE REFERENCE: FCSB-100-Div. 1
; CURRENT APPLICATION NUMBER: US/10/825,632
; CURRENT FILING DATE: 2004-04-15
; PRIOR APPLICATION NUMBER: US 10/070,464
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: PCT/AU00/01085
; PRIOR FILING DATE: 2000-09-11
; PRIOR APPLICATION NUMBER: AU PQ5709
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: AU PQ2762
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 882
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-825-632-1

Alignment Scores:
Pred. No.: 1,97e-239      Length: 882
Score: 2422.00      Matches: 465
Percent Similarity: 82.3%      Conservative: 0
Best Local Similarity: 82.3%      Mismatches: 0
Query Match: 81.9%      Indels: 100
DB: 4      Gaps: 1

```

QY 482 TCCAGTGGTGGCTGCTGCTCCCAAGTGATTTCAAGTGCTCTATCAAGAGGAGATAGCA 541
Db 478 SerSerGlyGlyLeuProAlaProSerAspPheLysCysProIleLysGluGluAla 497
QY 542 ATTACCAAGTGGTGAAGTTCTTGGCCGCGATGGATCTAATATCCAAAGTGTATGAA 601
Db 498 IleThrSerGlyGluTrpGluValLeuGlyA-GHisGlySerAsnIleGlnValAspGlu 517
QY 602 GTACAGAGGCTGATATTTGAGGCCACCAAGACTCCCTTTAGAGCATCAGCTGTAC 661
Db 518 ValArgGluValThrPheGluGlyThrLysAspSerProLeuGluHisLysLeuTyr 537
QY 662 GTAGTCAATTACGTAATCTCGAGAGGTGAACAAGCTGACTGACCGTGGCTACTACAT 721
Db 538 ValValSerTyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHis 557
QY 722 TCTTGCTGCATAGTCAGCACTGTGACTTCTTTTATAGTAAGTATAGTAACCAAGAAT 781
Db 558 SerCysCysIleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsn 577
QY 782 CCACACTGTGTGCTCTTTACAGCTATCAAGTCTCTGAAGATGACCCCAACTTGCAGAAC 841
Db 578 ProHisCysValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThr 597
QY 842 AAGCAATTTTGGCCCACTTTTGGATTGAGTTCAGCAGGTCTCTTCTGACTATATCTCTCA 901
Db 598 LysGluPheTrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProPro 617
QY 902 GAAATTTCTCTTTTGAAGTACTACTGATTTTACATTTATGGATGCTCTCAAGCCT 961
Db 618 GluIlePheSerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysPro 637
QY 962 CATGATCTACAGCTCGGAAGAAATATCTCTGCTGCTGTATCATATATGCTGCTCTCAG 1021
Db 638 HisAspLeuGlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGln 657
QY 1021 ----- 1021
Db 658 ValGlnLeuValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAla 677
QY 1021 ----- 1021
Db 678 SerLeuGlyTyrValValValValIleAspAsnArgGlySerCysHisArgGlyLeuLys 697
QY 1021 ----- 1021
Db 698 PheGluGlyAlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGly 717
QY 1021 ----- 1021
Db 718 LeuGlnTyrLeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGly 737
QY 1021 ----- 1021
Db 738 TrpSerTyrGlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArg 757
QY 1022 GTTGTATTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAA 1081
Db 758 ValAlaIleAlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGlu 777
QY 1082 CGTTATATGGTCAACCTGACAGAAATGAACAGGCTATTACTTAGATCTGTGGCCATG 1141
Db 778 ArgTyrMetGlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMet 797
QY 1142 CAAGCAGAAAGTTCCTCTGAAACCAATCGTTTACTGCTCTACATGCTTCTCGGAT 1201
Db 798 GlnAlaGluLysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAsp 817
QY 1202 GAGATGCTCAATTTGCAATACAGTATATTACTGAGTGTGTTTGTAGTGCGCTGGAAG 1261
Db 818 GluAsnValHisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLys 837

QY 1262 CCATATGATTTACAGATCTCTCTCAGGAGACACACAGCATTAAGAGTTCTCTGAATCGGA 1321
Db 838 ProTyrAspLeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGly 857
QY 1322 GAACATTATGAACGTGATCTTTTGCATCTACCTTCAAGAAACCTTGGATCAGTATTGCT 1381
Db 858 GluHisTyrGluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAla 877
QY 1382 GCTCTAAAAGTGATA 1396
Db 878 AlaLeuLysValIle 882
RESULT 9
US-10-982-512-1
; Sequence 1, Application US/10982512
; Publication No. US20050059081A1
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/10/982,512
; CURRENT FILING DATE: 2004-11-05
; PRIOR APPLICATION NUMBER: US/09/976,674
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 882
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-982-512-1
Alignment Scores:
Pred. No.: 1,97e-239 Length: 882
Score: 2422.00 Matches: 465
Percent Similarity: 82.3% Conservative: 0
Best Local Similarity: 82.3% Mismatches: 0
Query Match: 81.9% Indels: 100
DB: 5 Gaps: 1
US-10-825-632-6 (1-1669) x US-10-982-512-1 (1-882)
QY 2 ACAGGTACAGCAAAATCCTAAAGTCACCTTTTAAGATGTCAGAAATAATGATTGCTGAA 61
Db 318 ThrGlyThrAlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGlu 337
QY 62 GGAAGGATCATAGATGTCATAGATAAGGAACTAATTAACCTTTTGAGATTCTATTGAA 121
Db 338 GlyArgIleIleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGlu 357
QY 122 GGAGTTGAATATATTGCCAGAGCTGGATGGACTCTCTGAGGGGAAATATGCTTGTCCATC 181
Db 358 GlyValGluTyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIle 377
QY 182 CTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGATTTATTTATC 241
Db 378 LeuLeuAspArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIle 397
QY 242 CAGTAGACATGATCTTATGGAAAGGACAGACTCATTGATGAGTCCCTGATCTGTG 301
Db 398 ProValGluAspValMetGluArgGlnArgLeuIleGluSerValProAspSerVal 417
QY 302 ACGCCAATTAATCTATGAAGAAACAACAGACATCTGGATAAATATCCATGACATCTTT 361
Db 418 ThrProLeuIleGlyTyrGluIleThrThrAspIleTrpIleAsnIleHisAspIlePhe 437
QY 362 CATGTTTTCCTCCCAAGTCAGAGAGAAATTCAGTTTATTTTGTGCTCTGAATCGAA 421

Db	438	HisValPheProGlnSerHisGluGluGluLeuPheAlaSerGluCysLys	457	QY	1202	GAGATGTCATTTGCACATACAGTATATTACTGAGTTTTTTAGTGAGGCTGGAAG	1261
QY	422	ACAGGTTTCCGTCAATTTATATAAAATTTACATCTATTATTAAGGAAAGCAAAATATATAACGA	481	Db	818	GluAsnValHisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLys	837
Db	458	ThrGlyPheArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArg	477	QY	1262	CCATATGATTTACAGATCTATCTCTCAGGAGAGACACAGCATAGAGTTCTCTGAATCGGA	1321
QY	482	TCCAGTGTGGCTGCTGCTCCCAAGTGATTTCAAGTGTCCTTATCAAGAGAGATAGCA	541	Db	838	ProTyrAspLeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGly	857
Db	478	SerSerGlyGlyLeuProAlaProSerAspPheLysCysProIleLysGluIleAla	497	QY	1322	GAACATTATGAATGCGATCTTTTGCATCTACCTTCAGAAAACCTTCGATCACGATTATGCT	1381
QY	542	ATTACAGTGTGAATGGAGTCTTGGCGCGCATGGATCTAAATATCAAGTCTCATCAA	601	Db	858	GluHisTyrGluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAla	877
Db	498	IleThrSerGlyGluIleProGluValLeuGlyArgHisGlySerAsnIleGlnValAspGlu	517	QY	1382	GCTCTAAAAGTCGATA	1396
QY	602	GTCAGAAGCTGGTATATTTTGAAGGCCACCAAGAGATCTCCCTTTAGAGCATCACCTGTAC	661	Db	878	AlaLeuLysValIle	882
Db	518	ValArgArgLeuValTyrPheGluGlyThrLysAspSerProLeuGluHisIleLysTyr	537	RESULT 10			
QY	662	GTAGTCAGTTACGTAATCTCTGGAGAGGTGACAAAGCTGACCGTGGCTACTCAT	721	US-10-072-012-621			
Db	538	ValValSerTyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHis	557	; Sequence 621, Application US/10072012			
QY	722	TCTTCTGTCATCAGTCAGCATGTGACTTCTTTATAGTAAAGTATAGTAAACAGAGAT	781	; Publication NO. US2004003493A1			
Db	558	SerCysCysIleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsn	577	; GENERAL INFORMATION:			
QY	782	CGACACTGTGTCTCTTTACAGCTATCAAGTCTCAAGTCTCAAGTCAACCACTTGCAGAAACA	841	; APPLICANT: Tchernev, Velizar			
Db	578	ProHisCysValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThr	597	; APPLICANT: Spytek, Kimberly			
QY	842	RAGGAATTTGGCCACCATTTTGGATTTCAGACGCTCTCTCTGACTATATCTCTCTCA	901	; APPLICANT: Zernhusen, Bryan			
Db	598	LysGluPheTrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProPro	617	; APPLICANT: Patturajan, Meera			
QY	902	GAATTTTCTCTTTTGAAGTACTACTGGATTACATTTGATGGGATGCTCTACAGCT	961	; APPLICANT: Shinkets, Richard			
Db	618	GluIlePheSerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysPro	637	; APPLICANT: Li, Li			
QY	962	CATGATCTACAGCTGGAAGAAATATCCTACTGTGCTGTTCATATATATGTTGCTCTCAG	1021	; APPLICANT: Gangolli, Esha			
Db	638	HisAspLeuGlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGln	657	; APPLICANT: Fadigar, Muralidhara			
QY	1021	-----	1021	; APPLICANT: Anderson, David W.			
Db	658	ValGlnLeuValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAla	677	; APPLICANT: Rastelli, Luca			
QY	1021	-----	1021	; APPLICANT: Miller, Charles E.			
Db	678	SerLeuGlyTyrValValValIleAspAsnArgGlySerCysHisArgGlyLeuLys	697	; APPLICANT: Gerlach, Valerie			
QY	1021	-----	1021	; APPLICANT: Taupier Jr, Raymond J.			
Db	698	PheGluGlyAlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGly	717	; APPLICANT: Gusev, Vladimir Y.			
QY	1021	-----	1021	; APPLICANT: Colman, Steven D.			
Db	718	LeuGlnTyrLeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGly	737	; APPLICANT: Wolenc, Adam R.			
QY	1021	-----	1021	; APPLICANT: Pena, Carol E. A			
Db	738	TrpSerTyrGlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArg	757	; APPLICANT: Furtak, Katarzyna			
QY	1022	GTTCCTATTGCTGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACAGGAA	1081	; APPLICANT: Grosse, William M.			
Db	758	ValAlaIleAlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGlu	777	; APPLICANT: Alsobrook II, John P.			
QY	1082	CGTTATATGGTCACTCAGCAGAAATGAACAGGGCTATTACTTAGGATCTGTGGCCATG	1141	; APPLICANT: Lepley, Denise M.			
Db	778	ArgTyrMetGlyHisProAspGlnAsnGlnGlnTyrTyrLeuGlySerValAlaMet	797	; APPLICANT: Rieger, Daniel K.			
QY	1142	CAAGCAGAAAGTTCCCTCTGAACCAATCGTTTACTGCTCTTACATGGTTTCTCGAT	1201	; APPLICANT: Burgess, Catherine E.			
Db	798	GlnAlaGluLysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAsp	817	; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same			
				; FILE REFERENCE: 21402-258			
				; CURRENT APPLICATION NUMBER: US/10/072,012			
				; CURRENT FILING DATE: 2002-01-31			
				; PRIOR APPLICATION NUMBER: 60/265,102			
				; PRIOR FILING DATE: 2001-01-30			
				; PRIOR APPLICATION NUMBER: 60/265,514			
				; PRIOR FILING DATE: 2001-01-31			
				; PRIOR APPLICATION NUMBER: 60/265,517			
				; PRIOR FILING DATE: 2001-01-31			
				; PRIOR APPLICATION NUMBER: 60/265,412			
				; PRIOR FILING DATE: 2001-01-31			
				; PRIOR APPLICATION NUMBER: 60/265,395			
				; PRIOR FILING DATE: 2001-01-31			
				; PRIOR APPLICATION NUMBER: 60/266,406			
				; PRIOR FILING DATE: 2001-02-02			
				; PRIOR APPLICATION NUMBER: 60/266,767			
				; PRIOR FILING DATE: 2001-02-05			
				; PRIOR APPLICATION NUMBER: 60/267,057			
				; PRIOR FILING DATE: 2001-02-07			
				; PRIOR APPLICATION NUMBER: 60/266,975			
				; PRIOR FILING DATE: 2001-02-07			
				; PRIOR APPLICATION NUMBER: 60/267,459			
				; PRIOR FILING DATE: 2001-02-08			
				; Remaining Prior Application data removed - See File Wrapper or PALM.			
				; NUMBER OF SEQ ID NOS: 1391			
				; SOFTWARE: PatentIn Ver. 2.1			
				; SEQ ID NO 621			

```
; LENGTH: 883
; TYPE: PRT
; ORGANISM: Mus musculus
us-10-072-012-621

Alignment Scores:
Pred. No.: 1,496-229 Length: 883
Score: 2326.00 Matches: 443
Percent Similarity: 80.2% Conservative: 10
Best Local Similarity: 78.4% Mismatches: 12
Query Match: 78.7% Indels: 100
DB: 4 Gaps: 1

US-10-825-632-6 (1-1669) x US-10-072-012-621 (1-883)

QY 2 ACAGGTACAGCAATCTTAAGTCTACTTTTAAGATGTCAGAAATAATGATGATGCTGAA 61
DB 319 ThrGlyThrAlaAsnProLysValThrPheLysMetSerGluIleValValAspAlaAla 338
QY 62 GNAAGGATCATGATGTCATAGATAAGGAACATAATCAACCTTTTGAGATTTCTATTGAA 121
DB 339 GlyGlyIleIleAspValIleAspLysGluLeuValGlnProPheGluIleLeuPheGlu 358
QY 122 GGAGTTGAATATATTGCCAGAGCTGGATGGACTCTCTGAGGGAATAATGCTTGGTCCATC 181
DB 359 GlyValGluTyrIleAlaArgAlaGlyTrpThrProGluGlyLysHisAlaTrpSerIle 378
QY 182 CTACTAGATCGCTCCAGACTCGCTCACAGATAGTGTGATCTCACTGAATATTATTATC 241
DB 379 LeuLeuAspArgSerGlnThrHisLeuGlnIleValLeuIleSerProGluLeuPheIle 398
QY 242 CCAGTAGAAGATGATGTTATGGAAGGAGAGACTCATTGAGTCAGTGCCTGATCTGTG 301
DB 399 ProValGluAspAlaMetAspArgGlnArgLeuIleGluSerValProAspSerVal 418
QY 302 AGCCACATAATATCTATGAAGAACACAGACATCTGGAATATATATCATCATCATCTTT 361
DB 419 ThrProLeuIleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePhe 438
QY 362 CATGTTTTTCCCAAGTCAGAGAGAAATGAGTTATTTTTCCTCTCAATGCGAAA 421
DB 439 HisValPheProGlnThrHisGluAspGluIleGluPheIleAspSerGlyLys 458
QY 422 ACAGGTTTCCGCTCATTTATACAAATTTACATCTATTTTAAGGAAGCAATATAACGA 481
DB 459 ThrGlyPheArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTrpLysArg 478
QY 482 TCAGTGTGGTGCCTGCTCCAGTGATTTCAAGTGTCCTATCAAGAGAGAGATAGCA 541
DB 479 SerSerGlyGlyLeuProAlaProSerAspPheLysCysProIleLysGluIleThr 498
QY 542 ATTACAGTGTGAATGGGAAGTCTTTGGCCGCATGGATCTAATATCCAAAGTTGATGAA 601
DB 499 IleThrSerGlyGluTrpGluValLeuGlyArgHisGlySerAsnIleTrpValAspGlu 518
QY 602 CTCAGAGGCTGGTATATTTTGAAGGCCAACCAAGACTCCCTTTTAGAGCATCACCTGTAC 661
DB 519 AlaArgLysLeuValTyrPheGluGlyThrLysAspSerProLeuGluHisIleLeuTyr 538
QY 662 GTAGTCAGTACGTAAATCTCGAGAGGTGACAAAGGCTGACGCGGTGCTACTCATCAT 721
DB 539 ValThrSerTyrAlaAsnProGlyGluValValArgLeuThrAspArgGlyTyrSerHis 558
QY 722 TCTTGCTGCATCAGTCAGACTGTGACTTCTTTAAGTAAGTATAGTAACAGAGAAAT 781
DB 559 SerCysCysLeuSerArgHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsn 578
QY 782 CCACACTGTGTGCTCTTTACAAGCTATCAAGTCTGAGATGACCACTTGCAGAAACA 841
DB 579 ProHisCysValSerLeuTyrLysLeuSerSerProGluAspAspProValHisLysThr 598
QY 842 AAGGAATTTGGGCCACATTTTGGATTTCAGAGGCTCTCTCTCTGACTATACCTCTCCA 901
```

```
DB 599 LysGluPheTrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProPro 618
QY 902 GAAATTTTCTCTTTGAAAGTACTACTGATTTTACATTTGATGGATCTCTCAAGCCT 961
DB 619 GluIlePheSerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysPro 638
QY 962 CATGATCTACAGCTCGAAGAAATATCTACTGTGCTGTTTCATATATGGTGGTCTCAG 1021
DB 639 HisAspLeuGlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGln 658
QY 1021 ----- 1021
DB 659 ValGlnLeuValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAla 678
QY 1021 ----- 1021
DB 679 SerLeuGlyTyrValValValIleAspAsnArgGlySerCysHisArgGlyLeuLys 698
QY 1021 ----- 1021
DB 699 PheGluGlyAlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGly 718
QY 1021 ----- 1021
DB 719 LeuGlnTyrLeuAlaSerGlnTyrAspPheIleAspLeuAspArgValGlyIleHisGly 738
QY 1021 ----- 1021
DB 739 TrpSerTyrGlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArg 758
QY 1022 GTTCTATTGCTGGGGCCAGTCACTCTGPGGATCTTCTATGATACAGGATACACGGAA 1081
DB 759 ValAlaIleAlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGlu 778
QY 1082 CGTTATATGGGTCACCTTCAGCAATGAACAGGCTATTACTTAGGATCTGTGGCCATG 1141
DB 779 ArgTyrMetGlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMet 798
QY 1142 CAAGCAGAAAAAGTTCCTCTCTGAACCAAACTGTTTACTGCTCTTACATGTTTCTGGAT 1201
DB 799 GlnAlaGlyLysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAsp 818
QY 1202 GAGAATGTCCATTTTGACATACACAGTATATTACTGAGTGTGTTTTTAGTGGGCTGGAAG 1261
DB 819 GluAsnValHisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLys 838
QY 1262 CCATATGATTTACAGATCTATCTCTCAGGAGAGACACAGCATAAGAGTCTCTGAATCGGA 1321
DB 839 ProTyrAspLeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGly 858
QY 1322 GAACATTATGAATGCACTCTTTTGCACTACCTTCAAGAAAACCTTGGATCAGTATTGCT 1381
DB 859 GluHisTyrGluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAla 878
QY 1382 GCTCTAAAGTGATA 1396
DB 879 AlaLeuLysValIle 883

RESULT 11
US-10-275-505-2
; Sequence 2, Application US/10275505
; Publication No. US20040081961A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: DELEGANE, Angelo M.; LAL, Preeti G.
; APPLICANT: HAFALIA, April J.A.; PATTERSON, Chandra
; APPLICANT: WALIA, Narinder K.; KEARNEY, Liam
; APPLICANT: TRIBOULEY, Catherine M.; KHAN, Farrah A.
; APPLICANT: YAO, Monique G.; BAUGHN, Mariah R.
; APPLICANT: AZIMZAI, Yalda; ELLIOTT, Vicki S.
; APPLICANT: NGUYEN, Damiel B.; GANDHI, Ameena R.
; APPLICANT: YANG, Junning; HERNANDEZ, Roberto
; APPLICANT: POLICKY, Jennifer L.; LU, Dyung Aina M.
```

```
; APPLICANT: REDDY, Roopa M.; YUE, Henry
; APPLICANT: TANG, Y. Tom
; TITLE OF INVENTION: PROTEASES
; FILE REFERENCE: PI-0085 USN
; CURRENT APPLICATION NUMBER: US/10/275,505
; PRIOR FILING DATE: 2002-11-04
; PRIOR APPLICATION NUMBER: PCT/US01/14651
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: 60/209,402
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: 60/207,477
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/205,803
; PRIOR FILING DATE: 2000-05-17
; PRIOR APPLICATION NUMBER: 60/203,566
; PRIOR FILING DATE: 2000-05-11
; PRIOR APPLICATION NUMBER: 60/202,082
; PRIOR FILING DATE: 2000-05-04
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PERL Program
; SEQ ID NO 2
; LENGTH: 580
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 376067CD1
US-10-275-505-2

Alignment Scores:
Pred. No.: 3,02e-221 Length: 580
Score: 2244.50 Matches: 439
Percent Similarity: 84.0% Conservative: 3
Best Local Similarity: 83.5% Mismatches: 5
Query Match: 75.9% Indels: 79
DB: 4 Gaps: 6

US-10-825-632-6 (1-1669) x US-10-275-505-2 (1-580)

Qy 2 AAGGTACAGCAATCTTAAAGTCACTTTTAAAGTGCAGAAATAATGATGATGCTGAA 61
Db |||
73 ThrGlyThrAlaAsnProLysValThrPheLysMetSerGluLeuMetIleAspAlaGlu 92
Qy 62 GGAAGATCATAGATCATAGATAAGAACTAATTCACCTTTTGAGATCTATTGAA 121
Db |||
93 GlyArgIleIleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGlu 112
Qy 122 GGAAGTGAATATATCCAGAGCTGGATGGACTCCTGAGGAGAAATATGCTTGTCATC 181
Db |||
113 GlyValGluIleAlaAlaAlaGlyThrProGluGlyLysIleAlaIlePheSerIle 132
Qy 182 CTACTAGATCGCTCCAGACTCGCTCAGATAGTGTGATCTCACCTGAAATATTATC 241
Db |||
133 LeuLeuAspArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIle 152
Qy 242 CCAGTGAAGATGATTTATGAAAGGAGAGACTCAATGAGTCAGTCCCTGATCTGTG 301
Db |||
153 ProValGluAspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerVal 172
Qy 302 ACGCCACTAATTATCTATGAAGAAACAACAGACATCTGGATAAATATCCATGACATCTT 361
Db |||
173 ThrProLeuIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIle 192
Qy 362 CAGTGTTCCTCCCAAGTCCAGAGGAAATGAGTTATTTTGTCTCTGATGCATA 421
Db |||
193 HisValPheProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLys 212
Qy 422 ACAGGTTCCGTCATTTATACAAAATTACATCTATTTTAAAGGAAAGCAATATAACGA 481
Db |||
213 ThrGlyPheArgHisLeuIleIleIleIleIleIleIleIleIleIleIleIleIleIle 232
Qy 482 TCCAGTGGTGGGCTGCTCCCAAGTGATTTCAAGTGTCTCATCAAGAGGAGATAGCA 541
Db |||

233 SerSerGlyGlyLeuProAlaProThr 241
542 ATTACAGTGGTGAATGGGAAGTT 571
242 ValThr 571
572 -----CGGCATGGATCTAATATCCAAAGTTGATGAAGTCAGAAAGGCTGGTATATTTGAA 625
259 CysValThrHisIleValGluIleGlnValAspGluValArgLeuValThrPheGlu 278
626 GGCACCAAGACATCCCTTTTAGAGCATCACCTGTAGTAGTCAGTACGTAAATCCCTGGA 685
279 GlyThrIleAspSerProLeuGluHisIleValIleValIleValIleValIleValIle 298
686 GAGGTGAAGGCTGATGACCGTGGCTACTCACAATCTTCTGCTGATCAGTCAGCAGCTGT 745
299 GluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCysIleSerGlnHisCys 318
746 GACTTCTTTATAGTAAGTATAGTAACACAGAGAAATCCACACTGTGTGCTCCCTTTACAAG 805
319 AspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeuTyrLys 338
806 CTATCAAGTCTCAAGATGACCAACTGCAAAACAAAGGAATTTTGGCCACCATTTTG 865
339 LeuSerSerProGluAspAspProThrCysLysThrLysGluPheTrpAlaThrIleLeu 358
866 GATTACAGCAGTCTCTCTGCTGACTATATCTCTCCAGAAATTTCTCTTTTGAAGTACT 925
359 AspSerAlaGlyProLeuProAspTyrThrProProGluIlePheSerPheGluSerThr 378
926 ACTGGATTTACATTTGATGGATGCTCTACAAAGCTCATGATCTACAGCTGGGAAAGAAA 985
379 ThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGlyLysLys 398
986 TATCCTACTGCTGCTTCATATATATGTTGCTCTCAG----- 1021
399 TyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeuValAsnAsnArgPhe 418
1021 ----- 1021
419 LysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrValValVal 438
1021 ----- 1021
439 IleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLysTyrLys 458
1022 ---GTTGCTATTGCTGGGCCCCAGTCACCTCTGTGGATCTTCTATGATACAGGATACACG 1078
459 MetValAlaIleAlaGlyAlaProValThrLeuThrPheTyrAspThrGlyTyrThr 478
1079 GAACGTTATATGGTCACTGACCGACGACGAAATGAACAGGGCTATTACTTAGGATCTGTGGCC 1138
479 GluArgTyrMetGlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAla 498
1139 ATGCAAGCAGAAAGTTCCCTCTGACCAAAATCGTTTACTGCTCTTACATGTTCTCTG 1158
499 MetGlnAlaGluLysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeu 518
1199 GATCAGAAATCTCCATTTTGCATACATCATATATTACTGAGTTTATAGTCAGGCTGGA 1258
519 AspGlnAsnValHisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGly 538
1259 AAGCCATATCATTTACAGATCTCTCAGGAGACACAGCATAAGAGTTCTCTGAATCG 1318
539 LysProTyrAspLeu-----GlnGluArgHisSerIleArgValProGluSer 554
1319 GGAGAACATTATGAATGATCTTTTGTCTACCTTCAAGAAACCTTGGATCAGCTATT 1378
555 GlyGluHisTyrGluLeuHisLeuLeuHisTyrLeuGlnGlnAsnLeuGlySerArgIle 574
1379 GCTGCTCTAAAGTGATA 1396
575 AlaAlaLeuLysValIle 580
```

RESULT 12

```
US-11-140-224-2
; Sequence 2: Application US/111140224
; Publication No: US20050227280A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: DELEGANS, Angelo M.; LAL, Preeti G.
; APPLICANT: HAPALIA, April J.A.; PATTERSON, Chandra
; APPLICANT: WALIA, Narinder K.; KEARNEY, Liam
; APPLICANT: TRIBOULEY, Catherine M.; KHAN, Farrah A.
; APPLICANT: YAO, Monique G.; BAUGHN, Mariah R.
; APPLICANT: AZIMZAI, Yalda; ELLIOTT, Vicki S.
; APPLICANT: NGUYEN, Darniel B.; GANDHI, Ameena R.
; APPLICANT: YANG, Junming; HERNANDEZ, Roberto
; APPLICANT: POLICKY, Jennifer L.; LU, Dyung Aina M.
; APPLICANT: REDDY, Roopa M.; YUE, Henry
; APPLICANT: TANG, Y. Tom
; TITLE OF INVENTION: PROTEASES
; FILE REFERENCE: PI-0085 USN
; CURRENT APPLICATION NUMBER: US/11/140,224
; CURRENT FILING DATE: 2005-05-31
; PRIOR APPLICATION NUMBER: US/10/275,505
; PRIOR FILING DATE: 2002-11-04
; PRIOR APPLICATION NUMBER: PCT/US01/14651
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: 60/209,402
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: 60/207,477
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/205,803
; PRIOR FILING DATE: 2000-05-17
; PRIOR APPLICATION NUMBER: 60/203,566
; PRIOR FILING DATE: 2000-05-11
; PRIOR APPLICATION NUMBER: 60/202,082
; PRIOR FILING DATE: 2000-05-04
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PERL Program
; SEQ ID NO 2
; LENGTH: 580
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 376067CD1
US-11-140-224-2
Alignment Scores:
Pred. No.: 3,02e-221 Length: 580
Score: 2244.50 Matches: 439
Percent Similarity: 84.0% Conservative: 3
Best Local Similarity: 83.5% Mismatches: 5
Query Match: 75.9% Indels: 79
DB: 6 Gaps: 6
US-10-825-632-6 (1-1669) x US-11-140-224-2 (1-580)
Qy 2 ACAGGTACAGCAAACTCTTAAGTCACTTTTAAAGATGTCAGAAATAATGATGCTGAA 61
Db 73 ThrGlyThrAlaasnProlysvalThrPheLysMetSerGluileMetileaspAlaGlu 92
Qy 62 GGNAGCATCATAGTCTCATAGATAGCAACTATTCACCTTTTGAGATCTTATTGAA 121
Db 93 GlyArgileaspValileaspLysGluileleGlnProPheGluileleuPheGlu 112
Qy 122 GGNATTGAATATATTCCAGAGCTGGATGGACTCTCCCTGAGGGAAAAATATGCTTGGTCCATC 181
Db 113 GlyValGluTyrlleAlaArgAlaGlyTrpThrProGluGlyLysTyrlaTrpSerile 132
Qy 182 CTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACTGAATATTATTATC 241
Db 133 LeuLeuAspArgSerGlnThrArgLeuGlnileValLeuileSerProGluLeuPheile 152
```

```
Qy 242 CCAGTAGAAGATCATGCTTATGGAAGGCGAGAGACTCATTGAGTCACTGCTCCCTGATTCGTG 301
Db 153 ProValGluAspValMetGluArgGlnArgLeuileGluSerValProAspSerVal 172
Qy 302 AGCCACTAATTATCTATGAAGAAACAACAGACATCTCGATAATAATATCCATGACATCTTT 361
Db 173 ThrProLeuileileTyrlGluGluThrThrAspilleTrpilleAsnilleHieAspilePhe 192
Qy 362 CATGTTTTCCCAAAAGTCACGAAGAGAAATTTAGTTTATTTTTCCTCTCGAATGCAAA 421
Db 193 HieValPheProGlnSerHieGluGluGluileGluPheilePheAlaserGluCysLys 212
Qy 422 ACAGGTTCCGTCATTTATACAAAATTACATCTATTTTAAAGGAAAGCAAAATATAACGA 481
Db 213 ThrGlyPheArgHieLysTyrlLysilleThrSerilleLeuLysGluSerLysTyrlLysArg 232
Qy 482 TCCAGTGTGGGCTGCTCTCAAGTGATTTTCAAGTGTCTATCAAAAGAGGAGATAGCA 541
Db 233 SerSerGlyGlyLeuProAlaProThr
Qy 542 ATTACACAGTGGTGAATGGGAAGTT-----CTTGGC----- 571
Db 242 ValThr-----TrpMetileThrPheMetArgSerLeuGlyThrProSerCysMet 258
Qy 572 -----CGCATCGATCTAATATCAAGTTGATGAAGTCAGAAGCGCTGATATATTTGAA 625
Db 259 CysValThrHieileValGluileGlnValaspGluValArgLeuValTyrlPheGlu 278
Qy 626 GGCACCAAGAGCTCCCTTTAGAGCATCACCTGACGTAGTACGTAGTAAATCTCTGGA 685
Db 279 GlyThrLysAspSerProLeuGluHieLysleutyrlValSerTyrlValAsnProGly 298
Qy 686 GAGGTGACAAGGCTGACCCGCTGCTACTCACATTTCTGCTGATCAGTCAGTACTGT 745
Db 299 GluValThrArgLeuThrAspArgGlyTyrlSerHieSerCysLysleSerGlnHieCys 318
Qy 746 GACTTCTTTTATAAGTAAGTATAGTAACCAAGAGAAATCCACACTGTGTGCTCCCTTACAAG 805
Db 319 AspPhePheileSerLysTyrlSerAsnGlnLysAsnProHieCysValSerLeuTyrlLys 338
Qy 806 CTATCAAGTCCTGGAAGATGACCCAACTTGCAGAAACAAAGGAAATTTTGGGCCACCATTTG 865
Db 339 LeuSerSerProGluAspAspProThrCysLysThrLysGluPheTrpAlaThrileLeu 358
Qy 866 GATTCTAGCAGGTCCTCTTCTGACTATATGCTCCAGAAATTTTCTCTTTTGAAGTACT 925
Db 359 AspSerAlaGlyProLeuProAspTyrlThrProProGluilePheSerPheGluSerThr 378
Qy 926 ACTGGATTATACATTGATGGGATGCTTACAAGGCTCATGATCTACAGCCTGGAAGAA 985
Db 379 ThrGlyPheThrLeuTyrlGlyMetLeuTyrlLysProHieAspLeuGlnProGlyLysLys 398
Qy 986 TATCCTACTGTGCTGTTTATATATGCTGCTGCTCAG----- 1021
Db 399 TyrProThrValLeuPheileTyrlGlyGlyProGlnValGlnLeuValAsnAsnArgPhe 418
Qy 1021 ----- 1021
Db 419 LysGlyValLysTyrlPheArgLeuAsnThrLeuAlaserLeuGlyTyrlValValVal 438
Qy 1021 ----- 1021
Db 439 IleAspAsnArgGlySerCysHieSArgGlyLeuLysPheGluGlyAlaPheLysTyrlLys 458
Qy 1022 ---GTTGCTATTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACAG 1078
Db 459 MetValAlaileAlaGlyAlaProValThrLeuTrpillePheTyrlAspThrGlyTyrlThr 478
Qy 1079 GAAAGTTATATGGTCCACCTGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCC 1138
Db 479 GluArgTyrlMetGlyHieProAspGlnAsnGluGlnGlyTyrlTyrlLeuGlySerValAla 498
Qy 1139 ATGCAAGCAGAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGGTTCCTG 1198
```

Db	499	MetGlnAlaGluLysPheProSerGluProAsnArgLeuLeuLeuLeuHisGlyPheLeu	518
QY	1199	GATGAGAATGTCATTTTGCACATACAGTATATTAATCTAGCTTTTGTAGTGAGCGCTGGA	1258
Db	519	AspGluAsnValHisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGly	538
QY	1259	AAGCCATATGATTACAGATCTATCTCAGGAGAGACACAGCATAAAGATTCCTGAATCG	1318
Db	539	LysProTyrAspLeu-----GlnGluArgHisSerIleArgValProGluSer	554
QY	1319	GGAGAACTATGAACTGTCATCTTTTGCACCTACCTTCAAGAAACCTTGGATCAGCTATT	1378
Db	555	GlyGluHisTyrGluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIle	574
QY	1379	GCTGCTCTAAAGTGATA	1396
Db	575	AlaAlaLeuLysValIle	580
RESULT 13			
US-09-976-674-7			
; Sequence 7, Application US/09976674			
; Patent No. US20020115843A1			
; GENERAL INFORMATION:			
; APPLICANT: Qi, Steve			
; APPLICANT: Akinsanya, Karen			
; APPLICANT: Riviere, Pierre			
; APPLICANT: Junien, Jean-Louis			
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV			
; FILE REFERENCE: 70669			
; CURRENT APPLICATION NUMBER: US/09/976,674			
; CURRENT FILING DATE: 2001-10-12			
; PRIOR APPLICATION NUMBER: US 60/240,117			
; PRIOR FILING DATE: 2000-10-12			
; NUMBER OF SEQ ID NOS: 61			
; SOFTWARE: PatentIn version 3.1			
; SEQ ID NO 7			
; LENGTH: 690			
; TYPE: PRT			
; ORGANISM: Homo sapiens			
US-09-976-674-7			
Alignment Scores:			
Pred. No.:	1.1e-191	Length:	690
Score:	1957.50	Matches:	372
Percent Similarity:	99.5%	Conservative:	1
Best Local Similarity:	99.2%	Mismatches:	0
Query Match:	66.2%	Indels:	2
DB:	3	Gaps:	1
US-10-825-632-6 (1-1669) x US-09-976-674-7 (1-690)			
QY	2	ACAGGTACAGCAAACTCTAAGTCACTTTAAGATGTCAGAAATAATGATGCTGAA	61
Db	318	ThrGlyThrAlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGlu	337
QY	62	GHAGATCATAGATGTCATAGTAAGCACTAATTCACCTTTGAGATTCTATTGAA	121
Db	338	GlyArgIleAspValIleAspLysGluLeuIleGlnProPheGluLeuPheGlu	357
QY	122	GDAGTTGAATATATTCAGAGCTCGATGGAGTCTCTCAGGGAATAATGCTCGTCCATC	181
Db	358	GlyValGluTyrIleAlaArgAlaGlyThrProGluGlyLysTyrAlaTrpSerIle	377
QY	182	CTACTAGATCGCTCCAGACTCGCTCATAGATAGTGTGATCTCACTGAATTATTATC	241
Db	378	LeuLeuAspArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIle	397
QY	242	CCAGTACAGATGATTTATGAAAGGAGAGACTCATTTAGTCACTGCTGCTGATCTGTG	301
Db	398	ProValGluAspValMetGluArgGlnArgLeuIleGluSerValProAspSerVal	417
QY	302	ACGCCATTAATTATCTATGAAGAAACACAGACATCTGGATAAATATCCATGACATCTTT	361
Db	418	ThrProLeuIleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePhe	437
QY	362	CATGTTTTTCCCAAGTACAGAGAAATAGATTATTTTTCCTCTGAATGCAAA	421
Db	438	HisValPheProGlnSerHisGluGluIleGluPheIlePheAlaSerGluCysLys	457
QY	422	ACAGTTTTCGTATTATACAAATATACATCTATTTTAAAGAAAGCAAAATATATAACA	481
Db	458	ThrGlyPheArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArg	477
QY	482	TCCAGTGTGGGCTGCTGCTCCAAAGTATTTCAAGTGTCTCTATCAAGAGAGAGATACA	541
Db	478	SerSerGlyGlyLeuProAlaProSerAspPheLysCysProIleLysGluIleAla	497
QY	542	ATTACCACTGTGTAATGGGAAGTTCTTGGCCGCATCGATCTAATATCAAGTTGATGAA	601
Db	498	IleThrSerGlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGlu	517
QY	602	GTCAAGAGCTGTATATTTTGAAGCCACCAAGACTCCCTTTAGAGCATCACCTGTAC	661
Db	518	ValArgArgLeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeuTyr	537
QY	662	GTAGTCATGTTACGTAATCTCTGGAGAGGTGACAGGCTGACTGACCGTGGCTACTCACAT	721
Db	538	ValValSerTyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHis	557
QY	722	TCCTGCTGCATCAGTCAGCTGTGACTTCTTTATAGTAAGTATAGTAACAGAGAAAT	781
Db	558	SerCysCysIleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsn	577
QY	782	CCACACTGTGTCTCCCTTTTACAAGCTATCAAGTCTCTGAAGATGACCCAACTTGCAAAACA	841
Db	578	ProHisCysValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThr	597
QY	842	AAGGAATTTTGGGCCACCATTTTGGATTGACAGGCTCTCTCTGCTGCTATCTCTCCA	901
Db	598	LysGluPheTrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProPro	617
QY	902	GAATTTTCTCTTTGAAAGTACTACTGGATTATGATGGGATGCTGTACAAAGCCT	961
Db	618	GluIlePheSerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysPro	637
QY	962	CATGATCTACAGCTCGGAAAGAAATATCTTACTGTCTGTCTGTATATATATGCTGTCTAG	1021
Db	638	HisAspLeuGlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGly---Arg	656
QY	1022	GTTGCTATTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGATACACGAA	1081
Db	657	-LeuLeuLeuLeuGlyProGlnSerLeuCysGlySerSerMetIleGlnAspThrArgAs	676
QY	1082	CGTTATATGGTCCCTGACCAGAAATGACAGGGCTATTACT	1124
Db	676	nValIleTrpValThrLeuThrArgMetAsnArgAlaIleThr	690
RESULT 14			
US-10-982-512-7			
; Sequence 7, Application US/10982512			
; Publication No. US20050059081A1			
; GENERAL INFORMATION:			
; APPLICANT: Qi, Steve			
; APPLICANT: Akinsanya, Karen			
; APPLICANT: Riviere, Pierre			
; APPLICANT: Junien, Jean-Louis			
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV			
; FILE REFERENCE: 70669			
; CURRENT APPLICATION NUMBER: US/10/982,512			
; CURRENT FILING DATE: 2004-11-05			
; PRIOR APPLICATION NUMBER: US/09/976,674			
; PRIOR FILING DATE: 2001-10-12			
; PRIOR APPLICATION NUMBER: US 60/240,117			
; PRIOR FILING DATE: 2000-10-12			
; NUMBER OF SEQ ID NOS: 61			

```
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 690
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-982-512-7

Alignment Scores:
Pred. No.: 1,1e-191 Length: 690
Score: 1957.50 Matches: 372
Percent Similarity: 99.5% Conservative: 1
Best Local Similarity: 99.2% Mismatches: 0
Query Match: 66.2% Indels: 2
DB: 5 Gaps: 1

US-10-825-632-6 (1-1669) x US-10-982-512-7 (1-690)

QY 2 ACAGGTACAGCAATCTTAAGTCACTTTTAAGATGTCAGAAATATGATGCTGAA 61
Db 318 ThrGlyThrAlaAsnProLysValThrPheLysMetSerGluLeuMetileAspAlaGlu 337
QY 62 GGNAGGCATAGATGTCATAGATAAGGAACCTAATTCACCTTTTGGATTTCTATTGAA 121
Db 338 GlyArgileAspValileAspLysGluLeuileGlnProPheGluileLeuPheGlu 357
QY 122 GCAGTTGAATATTTGCCAGAGCTGGATGGACTCCTGAGGGAATATGCTTTGGTCCATC 181
Db 358 GlyValGluTyrlleAlaArgAlaGlyTrpThrProGluGlyLysTyAlaTrpSerile 377
QY 182 CTACTAGATCGTCCAGACTCGCTACAGATAGTGTGATCTCACTCACTGAATATTATTC 241
Db 378 LeuLeuAspArgSerGlnThrArgLeuGlnileValleuileSerProGluLeuPheile 397
QY 242 CCAGTAGAAGATGATGTTTGGAAAGGCAGAGACTCATTGAGTCAGTGCCTGATTCCTG 301
Db 398 ProValGluAspAspValMetGluArgGlnArgLeuileGluSerValProAspSerVal 417
QY 302 ACGCCTCAATATCTATGAAAGAAACACAGACATCTGGATAAATATCCATGACATCTTT 361
Db 418 ThrProLeuileileTyrlleGluGluThrThrAspilleTrpilleAsnileHisAspillePhe 437
QY 362 CATGTTTTTCCCAAGTACGAGAGAGAAATGAGTTTATTTTGGCTCTGAAATGCAAA 421
Db 438 HisValPheProGlnSerHisGluGluGluPheilePheAlaSerGluCysAlys 457
QY 422 ACAGGTTTCCGTCAATTTATACAAATATACATCTATTTAAAGGAAGCAATATAACGA 481
Db 458 ThrGlyPheArgHisLeuTyrlleThrSerilleLeuLysGluSerileTyrlsArg 477
QY 482 TCCAGTGGTGGCTGCTCCAGAGTATTCAGTGATTCAGTGCTCCTATCAAGAGAGATAGCA 541
Db 478 SerSerGlyGlyLeuProAlaProSerAspPheLysCysProilleLysGluGluileAla 497
QY 542 ATTACCAAGTGGTGAATGGNAGTCTTGGCCGGCATGGATCTATATATCCAGTTGATGAA 601
Db 498 IleThrSerGlyLutrpGluValleuGlyArgHisGlySerAsnilleGlnValAspGlu 517
QY 602 GTCAGAGGCTGGTATATTTTGAAGCACCAAGACTCCCTTTTAGAGCATCACCTGTATC 661
Db 518 ValArgArgLeuValTyrlleGluGlyThrLysAspSerProLeuGluHisHisLeuTyrl 537
QY 662 GTAGTCAGTTACTAAATCCTGGAGAGGTGACAAGGCTGACTACCGTGGCTACTACAT 721
Db 538 ValValSerTyrlleAsnProGlyLutrpGluValThrArgLeuThrAspArgGlyTyrlSerHis 557
QY 722 TCTTGCTGATCAGTCAGTGTGACTTCTTTTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 781
Db 558 SerCysCysileSerGlnHisCysAspPhePheileSerLysTyrlSerAsnGlnLysAsn 577
QY 782 CCACACTGTGTGCTCTTTTACAAAGCTATCAAGTCTCTGAAGATGACCAACTTCCAAACAA 841
Db 578 ProHisCysValSerLeuTyrlLysLeuSerSerProGluAspAspProThrCysLysThr 597
```

```
QY 842 AAGGAATTTGGGCCACCACTTTTGGATTACAGCAGTCTCTTCTCTGACTATCTCTCCA 901
Db 598 LysGluPheTrpAlaThrileLeuAspSerAlaGlyProLeuProAspTyrlThrProPro 617
QY 902 GAAATTTCTCTTTTGAAGTACTACTGGATTACATTGTATGGGATGCTCTCAACAGCCT 961
Db 618 GluilePheSerPheGluSerThrThrGlyPheThrLeuTyrlGlyMetLeuTyrlLysPro 637
QY 962 CATGATCTACAGCCTGGAAAGAAATATCTACTGTGCTGTGTTCATATATGGTGGTCTCAG 1021
Db 638 HisAspLeuGlnProGlyLysLysTyrlProThrValLeuPheileTyrlGlyGly---Arg 656
QY 1022 GTTGCTATTGCTGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGAA 1081
Db 657 -LeuLeuLeuLeuGlyProGlnSerLeuCysGlySerSerMetileGlnAspThrArgAs 676
QY 1082 CGTTATATGGGTCACTCCCTGACCAGAAATGAACAGGCTATTACT 1124
Db 676 nValilleTrpValThrLeuThrArgMetAsnArgAlaileThr 690

RESULT 15
US-09-976-674-19
; Sequence 19, Application US/09976674
; Patent No. US20020115843A1
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akisanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 0869
; CURRENT APPLICATION NUMBER: US/09/976,674
; PRIORITY FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIORITY FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19
; LENGTH: 658
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-19

Alignment Scores:
Pred. No.: 2,59e-176 Length: 658
Score: 1808.00 Matches: 338
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 61.1% Indels: 0
DB: 3 Gaps: 0

US-10-825-632-6 (1-1669) x US-09-976-674-19 (1-658)

QY 2 ACAGGTACAGCAATCTTAAGTCACTTTTAAGATGTCAGAAATATGATGATGCTGAA 61
Db 318 ThrGlyThrAlaAsnProLysValThrPheLysMetSerGluLeuMetileAspAlaGlu 337
QY 62 GGAAGGATCATAGATGTCATAGATAAGGAACCTAATTCACCTTTGAGATCTTATTGAA 121
Db 338 GlyArgileileAspValileAspLysGluLeuileGlnProPheGluileLeuPheGlu 357
QY 122 GGAGTTGAATATTCGAGAGCTGGATGGACTCCTGAGGGAATATGCTTTGGTCCATC 181
Db 358 GlyValGluTyrlleAlaArgAlaGlyTrpThrProGluGlyLysTyrlAlaTrpSerile 377
QY 182 CTACTAGATCGTCCAGACTCGCTACAGATAGTGTGATCTCACTCACTGAATATTATTC 241
Db 378 LeuLeuAspArgSerGlnThrArgLeuGlnileValleuileSerProGluLeuPheile 397
QY 242 CCAGTAGAAGATGATGTTTGGAAAGGCAGAGACTCATTGAGTCAGTGCCTGATTCCTG 301
Db 398 ProValGluAspAspValMetGluArgGlnArgLeuileGluSerValProAspSerVal 417
```

QY 302 AGCCCACTAATTTATCTATGAAGAAACAACAGACATCTGGATAAATATCCATGACATCTTT 361
 Db |||||
 418 ThrProLeuIleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePhe 437
 QY 362 CATGTTTCCCAAGTCACGAGAGAAATTCAGTTTATTTTGGCTCTGAATCATAA 421
 Db |||||
 438 HisValPheProGlnSerHisGluGluIleGluPheIlePheAlaSerGluCysLys 457
 QY 422 ACAGGTTTCCGTCATTATACAAAATTCATCTATTTAAAGGAAGCAAAATATAACGA 481
 Db |||||
 458 ThrGlyPheArgHisLeuTyrIleThrSerIleLeuLysGluSerLysTyrLysArg 477
 QY 482 TCCAGTGTGGCTGCTGCCAAGTCATTTCGAAGTCTCCTATCAAAAGAGAGATAGCA 541
 Db |||||
 478 SerSerGlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAla 497
 QY 542 ATTACAGTGGTGAATGGGAAGTCTTGGCCGGCATGGATCTAATATCCAAAGTTGATGA 601
 Db |||||
 498 IleThrSerGlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGlu 517
 QY 602 GTCAGAAAGCTGCTATATTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTAT 661
 Db |||||
 518 ValArgArgLeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeuTyr 537
 QY 662 GTAGTCAGTTACGTAATCCTGGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACAT 721
 Db |||||
 538 ValValSerTyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHis 557
 QY 722 TCTTGCTGCATCAGTCAGCAGCTGCTACTTCTTTAAGTAAGTATAGTACCAGAGAAT 781
 Db |||||
 558 SerCysCysIleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsn 577
 QY 782 CGACACTGTGTGCTCTTTACAAAGCTATCAAGTCTGAAGATGACCCAACTTGCAAAACA 841
 Db |||||
 578 ProHisCysValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThr 597
 QY 842 AAGGAATTTGGGCCACCAATTTTGATTTCAGCAGGCTCTCTTCTGACTATCTCTCCA 901
 Db |||||
 598 LysGluPheTrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProPro 617
 QY 902 GAAATTTCTCTTTGAAAGTACTACTGATTTACATTTGATGGGATGCTCTACAGCCT 961
 Db |||||
 618 GluIlePheSerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysPro 637
 QY 962 CATGATCTACAGCTGGAAGAAATATCTACTGTGCTGTTCATATATGCTGGT 1015
 Db |||||
 638 HisAspLeuGlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGly 655

Search completed: May 2, 2006, 03:08:15
 Job time : 277.002 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: May 2, 2006, 02:50:19 ; Search time 5:57883 Seconds
(without alignments)
2721.092 Million cell updates/sec

Title: US-10-825-632-6
Perfect score: 2957
Sequence: 1 aacaggtacagcaatccta.....aaaaaaaaaaaaaaaaaa 1669

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 232119 seqs, 45477862 residues

Total number of hits satisfying chosen parameters: 464238

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp
-Q=/abss/ABSSWEB_spool/US10825632/runat_01052006_110000_3533/app_query.fasta.1
-DB=Published Applications_AA_New -QPM=fastan -SUFFIX=n2p.rapbn -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsum62
-THRS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -HOST=abss02p
-USER=US10825632 @CGN_1.1.59 runat_01052006_110000_3533 -NCPU=6 -ICPU=3
-NO_WMAP -NEG SCORES=0 -WAIT -DSPELLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=10 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications_AA_New.*

1:	/SID55/ptodata/1/pubpaa/US08_NEW_PUB.pepl.*
2:	/SID55/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
3:	/SID55/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
4:	/SID55/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
5:	/SID55/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
6:	/SID55/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
7:	/SID55/ptodata/1/pubpaa/US09_NEW_PUB.pepl.*
8:	/SID55/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
9:	/SID55/ptodata/1/pubpaa/US10_NEW_PUB.pepl.*
10:	/SID55/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
11:	/SID55/ptodata/1/pubpaa/US11_NEW_PUB.pepl.*
12:	/SID55/ptodata/1/pubpaa/US60_NEW_PUB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	2422	81.9	882	11	US-11-151-601-20
2	1298.5	43.9	497	11	US-11-264-096-1594
3	1294.5	43.8	497	11	US-11-264-096-1593
4	588.5	19.9	349	11	US-11-264-096-1591
5	577	19.5	109	11	US-11-176-951-10

6	394	13.3	99	11	US-11-176-951-11
7	359.5	12.2	745	11	US-11-079-463-6408
8	269.5	9.1	627	11	US-11-079-463-7758
9	238	8.0	738	11	US-11-208-288-4
10	238	8.0	766	9	US-10-501-035-234
11	238	8.0	766	11	US-11-208-288-2
12	234	7.9	762	11	US-11-116-939-13
13	234	7.9	766	9	US-10-522-789-2
14	229	7.7	760	11	US-11-208-288-6
15	214	7.2	760	11	US-11-186-284-55
16	145	4.9	102	11	US-11-176-951-7
17	139	4.7	115	11	US-11-176-951-9
18	139	4.7	115	11	US-11-176-951-12
19	133.5	4.5	102	11	US-11-176-951-8
20	111.5	3.8	1019	9	US-10-995-561-982
21	109.5	3.7	959	9	US-10-467-962B-4
22	104.5	3.5	352	11	US-11-098-686-11006
23	104	3.5	468	11	US-11-156-084-250
24	102	3.4	732	9	US-10-518-599-22
25	98.5	3.3	623	11	US-11-188-298-16754
26	97	3.3	624	11	US-11-079-463-7504
27	95.5	3.2	570	11	US-11-143-984A-10
28	95.5	3.2	624	11	US-11-079-463-10187
29	94.5	3.2	2323	9	US-10-793-626-760
30	93	3.1	737	11	US-11-079-463-9281
31	92.5	3.1	421	11	US-11-045-004-2837
32	92.5	3.1	595	11	US-11-079-463-8616
33	92.5	3.1	2725	11	US-11-113-424-52
34	91.5	3.1	1461	11	US-11-183-136-28
35	91	3.1	439	11	US-11-075-185-11
36	90.5	3.1	328	11	US-11-188-298-20669
37	90.5	3.1	487	9	US-10-467-657-2272
38	89.5	3.0	486	11	US-11-188-298-21031
39	89.5	3.0	3580	9	US-10-510-941-14
40	89	3.0	563	11	US-11-079-463-9616
41	89	3.0	622	11	US-11-188-298-15979
42	89	3.0	628	11	US-11-087-099-4659
43	89	3.0	828	9	US-10-995-561-983
44	89	3.0	918	9	US-10-995-561-981
45	88.5	3.0	422	9	US-10-517-939-238

ALIGNMENTS

RESULT 1
US-11-151-601-20
; Sequence 20, Application US/11151601
; Publication No. US20060003413A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Olandt, Peter J.
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Curtis, Rory A. J.
; APPLICANT: Williamson, Mark
; APPLICANT: Weich, Nadine
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASE, PHOSPHATASE, AND PROTEASE FAMILY MEMBERS AND USES THEREOF
; FILE REFERENCE: MF100-054P/PC10MIDVIM
; CURRENT APPLICATION NUMBER: US/11/151,601
; CURRENT FILING DATE: 2005-06-13
; PRIOR APPLICATION NUMBER: US 10/170,789
; PRIOR FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: US 09/797,039
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: PCT/US01/06525
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/186,061
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 09/882,166
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: PCT/US01/19269
; PRIOR FILING DATE: 2001-06-15

; PRIOR APPLICATION NUMBER: US 60/212,078									
; PRIOR FILING DATE: 2000-06-15									
; PRIOR APPLICATION NUMBER: US 09/934,406									
; PRIOR FILING DATE: 2001-08-21									
; PRIOR APPLICATION NUMBER: PCT/US01/26052									
; PRIOR FILING DATE: 2001-08-21									
; PRIOR APPLICATION NUMBER: US 60/226,740									
; PRIOR FILING DATE: 2000-08-21									
; Remaining Prior Application data removed - See File Wrapper or PALM.									
; NUMBER OF SEQ ID NOS: 45									
; SOFTWARE: FastSeq for Windows Version 4.0									
; SEQ ID NO 20									
; LENGTH: 882									
; TYPE: PRT									
; ORGANISM: Homo sapiens									
US-11-151-601-20									
Alignment Scores:									
Pred. No.: 1.17e-225 Length: 882									
Score: 2422.00 Matches: 465									
Percent Similarity: 82.3% Conservative: 0									
Best Local Similarity: 82.3% Mismatch: 0									
Query Match: 81.9% Indels: 100									
DB: 11 Gaps: 1									
US-10-825-632-6 (1-1669) x US-11-151-601-20 (1-882)									
QY	2	ACAGTACAGCAAACTCTAAAGTCACCTTTAAGATGTCAGAAATAATGATTGATGCTGAA	61						
DB	318	ThrGlyThrAlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGlu	337						
QY	62	GGAAGGATCATAGATGTCATAGATAAGGAACCTAATCAACCTTTTGAGATTCATTTGAA	121						
DB	338	GlyArgIleIleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGlu	357						
QY	122	GGAGTTGAATATATGCCAGAGCTGGATGGACTCCTCGAGGAAATATGCTGGTCCATC	181						
DB	358	GlyValGluTyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIle	377						
QY	182	CTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGTGATCTCACCTGAAATATTATTC	241						
DB	378	LeuLeuAspArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIle	397						
QY	242	CCAGTAGAAGATGATGTTATGGAAGGACAGAGACTCATTGAGTCAGTGCCTGATTCGTG	301						
DB	398	ProValGluAspValMetGluArgGlnArgLeuIleGluSerValProAspSerVal	417						
QY	302	ACGCCACTAATTATCTATGAAGAAACAACAGACATCTGGATAAATATCCATGACATCTTT	361						
DB	418	ThrProLeuIleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePhe	437						
QY	362	CATGTTTTTCCCAAGTCACGAGAGAAATTCAGTTTTATTTTGGCTCTGAAATCAAA	421						
DB	438	HisValPheProGlnSerHisGluGluIleGluPheIlePheAlaSerGluCysLys	457						
QY	422	ACAGTTTCCGTCAATTTATCAAAATTACATCTATTTTAAAGGAAGCAATATAACGA	481						
DB	458	ThrGlyPheArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArg	477						
QY	482	TCBAGTGGTGGCTGCTCTCCAAAGTGATTTCAAGTCTCCTATCAAAAGGAGATAGCA	541						
DB	478	SerSerGlyGlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAla	497						
QY	542	ATTACCAAGTGGTAATGGAAAGTTCTTGGCCGGCATGGATCTAATATCCAAGTTGATGA	601						
DB	498	IleThrSerGlyGluTrpGluValLeuGluArgHisGlySerAsnIleGlnValAspGlu	517						
QY	602	GTCAAGAGGCTGGTATATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACCTGTAC	661						
DB	518	ValArgArgLeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeuTyr	537						
QY	662	GTAGTCAGTTACGTAATAATCCTCGAGAGGTGACAAGCTGACTGACCGTGGCTACTCACAT	721						

538	ValValSerTyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHis	557
722	TCITGCTGCATCAGTCAGCACTGACTTCTTTATAGTAAGTATATAGTACCAAGAAAT	781
558	SerCysCysIleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsn	577
782	CCACACTGTGTGTCCTTTTACAAGCTATCAAGTCTCGAAGATGACCAACTTGCAAAACA	841
578	ProHisCysValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThr	597
842	AAGGAATTTTGGGCCACCATTTTGGATTTCAGCAGGTCCTCTTCCTCACTATATCTCTCCA	901
598	LysGluPheTrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProPro	617
902	GAATTTTCTCTTTTGAAGTACTACTCGGATTTTACATTTGTATCGGATGCTCTCAACAGCCT	961
618	GluIlePheSerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysPro	637
962	CATGATCTACAGCCCTGGAAGAAATATCCTACCTGTGCTGTTCATATATGCTGCTCTCAG	1021
638	HisAspLeuGlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGln	657
1021	-----	1021
658	ValGlnLeuValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAla	677
1021	-----	1021
678	SerLeuGlyTyrValValValValIleAspAsnArgGlySerCysHisArgGlyLeuLys	697
1021	-----	1021
698	PheGluGlyAlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGly	717
1021	-----	1021
718	LeuGlnTyrLeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGly	737
1021	-----	1021
738	TrpSerTyrGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArg	757
1022	GTTGCTATTGCTGGGGCCCGAGTCACCTCTGTGTGATCTTCTATGATACAGGATACACGAA	1081
758	ValAlaIleAlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGlu	777
1082	CGTTATATGGGTCAACCTGACCAAGAATGAACAGGGCTATTACTAGGATCTGTGGCCATG	1141
778	ArgTyrMetGlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMet	797
1142	CAACGACAGAAAGTTCCCTCTCGAACCAATCGTTTACTGCTCTTACATGTTTCCCTGGAT	1201
798	GlnAlaGluLysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAsp	817
1202	GAGAACTCCATTTTGCACATACCAGTATATTACTCAGTTTTTTTACTGAGGGCTGAAAG	1261
818	GluAsnValHisPheAlaHisThrSerIleLeuLeuSerPheLeuValargAlaGlyLys	837
1262	CCATATGATTTACAGATCTATCTCTCAGAGAGACACAGCATAGAGTCTCTGATCGGCA	1321
838	ProTyrAspLeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGly	857
1322	GAACTATTGAACCTGCATCTTTTGCACCTACCTTCAAGAAAACCTTGGATCATCGTATTGCT	1381
858	GluHisTyrGluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAla	877
1382	GCTCTAAAGTGATA 1396	
878	AlaLeuLysValIle 882	
RESULT 2		
US-11-264-096-1594		
; Sequence 1594, Application US/11264096		

RESULT 2
US-11-264-096-1594
; Sequence 1594, Application US/11264096

Db	538	ValValSerTyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHis	557
QY	722	TCTTGCTGCATCAGTCAGCACTGTGACTTCTTTTATAAGTAAGTAAAGTAAACAGAA	781
Db	558	SerCysCysIleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsn	577
QY	782	CCACACTGTGTGCTCCCTTTACAAAGCTATCAAGTCTCTGAAGATGACCCCACTGCA	841
Db	578	ProHisCysValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThr	597
QY	842	AAGGAATTTTGGGCCACCACTTTTGGATTTCAGCAGGTCTCTTCTCTGACTATACCT	901
Db	598	LysGluPheTrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProPro	617
QY	902	GAATTTTCTCTTTTGAAGTACTACTGGATTACATTGATGGGATCTCTACAAAGCCT	961
Db	618	GluIlePheSerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysPro	637
QY	962	CATGATCTACAGCTCGAAAGAAATATCTACTGTCTGCTTTCATATATGGTGGTCTC	1021
Db	638	HisAspLeuGlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGln	657
QY	1021	-----	1021
Db	658	ValGlnLeuValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAla	677
QY	1021	-----	1021
Db	678	SerLeuGlyTyrValValValValIleAspAsnArgGlySerCysHisArgGlyLeuLys	697
QY	1021	-----	1021
Db	698	PheGluGlyAlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGly	717
QY	1021	-----	1021
Db	718	LeuGlnTyrLeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGly	737
QY	1021	-----	1021
Db	738	TrpSerTyrGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArg	757
QY	1022	GTTGCTATTCTGGGCCCCAGTCACCTCTGTGATCTTCTATGATACAGGATACACGAA	1081
Db	758	ValAlaIleAlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGlu	777
QY	1082	CGTTATATGGTCCACCTGACCAGAAATGAACAGGGCTATTACTTAGGATCTGTGCCATG	1141
Db	778	ArgTyrMetGlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMet	797
QY	1142	CAAGCAGAAAGTTCCCTCTGAACCAATCGTTTACTGCTCTTACATGGTTTCTCGAT	1201
Db	798	GlnAlaGluLysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAsp	817
QY	1202	GAGNATGCTCCATTTTCACATACCACTATATTTACTGAGTCTTTTAGTGAGGCTGGAAG	1261
Db	818	GluAsnValHisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLys	837
QY	1262	CCATATGATTTACAGATCTATCTCTCAGGAGACACAGCATAAAGTTCCTGAATCGGA	1321
Db	838	ProTyrAspLeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGly	857
QY	1322	GAACATTATGAATCGATCTTTTGGCATCTACCTTCAAGAAACCTTGGATACAGTATGCT	1381
Db	858	GluHisTyrGluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAla	877
QY	1382	GCTCTAAAGTGATA 1396	
Db	878	AlaLeuLysValIle 882	

```
; Publication No. US20060084794A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF546D1
; CURRENT APPLICATION NUMBER: US/11/264,096
; PRIORITY FILING DATE: 2005-11-02
; PRIOR APPLICATION NUMBER: 09/833,245
; PRIORITY FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIORITY FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIORITY FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIORITY FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 2267
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1594
; LENGTH: 497
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-264-096-1594

Alignment Scores:
Pred. No.: 3,91e-117 Length: 497
Score: 1298.50 Matches: 248
Percent Similarity: 61.4% Conservative: 57
Best Local Similarity: 49.9% Mismatches: 91
Query Match: 43.9% Indels: 101
DB: 11 Gaps: 2

US-10-825-632-6 (1-1669) x US-11-264-096-1594 (1-497)

QY 179 ATCTACTAGATCGCTCCAGACTCGGCTACAGATAGTGTGATCTCACTGAATATT 238
Db 1 MetPheLeuAspArgProGlnGlnTrpLeuGlnLeuValLeuLeuProAlaLeuPhe 20
QY 239 ATCCAGTAGAAGATGATGTTATGGAAGGAGAGACTATTGAGTCAGTGCCTGATCT 298
Db 21 IleProSerThrGluAsnGluGluGlnArgLeuAlaSerAlaArgAlaValProArgAsn 40
QY 299 GTGACGCCCTAATTAATCTATGAAGAACAACACACATCTGGATAAATATCCATGATC 358
Db 41 ValGlnProTyrValValTyrGluGluValThrAsnValTrpIleAsnValHisAspIle 60
QY 359 TTTTCATGTTTTCCCAAGTCAC---GAAGAGAAATTCAGTTTATTTTGCTCTGAA 415
Db 61 PheTyrProPheProGlnSerGluGluAspGluLeuCysPheLeuArgAlaAsnGlu 80
QY 416 TDCAAACAGGTTCCGTCATTATATACAAATTCATCTATTTAAAGMAAGCAATAT 475
Db 81 CysLeuThrGlyPheCysHisLeuTyrIysValThrAlaValLeuLeuLysSerGlnGlyTyr 100
QY 476 AACAGATCCAGTGGTGGGCTGCTCTCAAGTGAATTCAGTGTCTCCTATCAAGAGGAG 535
Db 101 AspTrpSerGluProPheSerProGlyGluAspGluPheLysCysProIleLysGluGlu 120
QY 536 ATAGCAATACAGTGGTGAATGGAAGTTCTTGGCCGGCATGATGATCAATATCAAGTT 595
Db 121 IleAlaLeuThrSerGlyGlnTrpGluValLeuAlaArgHisGlySerLysIleTrpVal 140
QY 596 GATGAAGTCAGAGGCTGGTATATTTTGAAGGACCAAGACTCCCTTTAGAGCATCAC 655
Db 141 AsnGluGluThrLysLeuValTyrPheGlnGlyThrLysAspThrProLeuGluHisHis 160
QY 656 CTGTACCTAGTACGTAATCTCTGAGAGGTCACAGGCTGACTGACCGTGGCTAC 715
Db 161 LeuTyrValValSerTyrGluAlaLaglyGluIleValArgLeuThrThrProGlyPhe 180
QY 716 TCATTTCTTGTGTCATCAGTACGACTGTGCTCTTTTATAAGTAAAGTATAGTAACAG 775
Db 181 SerHisSerCysSerMetSerGlnAsnPheAspMetPheValSerHisTyrSerSerVal 200

776 AAGAATCCACACTGTGTGTCCTCTTTAAGCTATATCAAGTCTGAAGATGACCCCACTTGC 835
Db 201 SerThrProProCysValHisValTyrLysLeuSerGlyProAspAspProLeuHis 220
QY 836 AAAACAAAGGAATTTTGGGCCACCATTTGGATTACAGAGTCCCTCTCTCCCTACTACT 895
Db 221 LysGlnProArgPheTrpAlaSerMetMetGluAlaAlaSerCysProProAspTyrVal 240
QY 896 CCTCCAGAAATTTCTCTTTGAAAGTACTACTGGATTTTACATTTGATGGGATGCTCTAC 955
Db 241 ProProGluIlePhePheHisThrAspSerAspValArgLeuTyrGlyMetIleTyr 260
QY 956 AAGCCTCATGATCTACAGCCTCGGAAGAAATATATCTCTGCTGCTGTCTATATATGTTGT 1015
Db 261 LysProHisAlaLeuGlnProGlyLysHisProThrValLeuPheValTyrGlyGly 280
QY 1016 CCTCAG----- 1021
Db 281 ProGlnValGlnLeuValAsnAsnSerPheLysGlyIleLysTyrLeuArgLeuAsnThr 300
QY 1021 ----- 1021
Db 301 LeuAlaSerLeuGlyTyrAlaValValIleAspGlyArgGlySerCysGlnArgGly 320
QY 1021 ----- 1021
Db 321 LeuArgPheGluGlyAlaLeuLysAsnGlnMetGlyGlnValGluIleGluAspGlnVal 340
QY 1021 ----- 1021
Db 341 GluGlyLeuGlnPheValAlaGluLysTyrGlyPheIleAspLeuSerArgValAlaIle 360
QY 1021 ----- 1021
Db 361 HisGlyTrpSerTyrGlyGlyPheLeuSerLeuMetGlyLeuIleHisLysProGlnVal 380
QY 1022 -----GTTGCTATTGCTGGGCCCCAGTCACTCTGTGGATCTCTTATGATACAGATAC 1075
Db 381 PheLysValAlaIleAlaGlyAlaProValThrValTrpMetAlaTyrAspThrGlyTyr 400
QY 1076 ACGAAACGTTATATGGGTACCTCCACCAAGATCAACAGGCTATTATTAGGATCTGTG 1135
Db 401 ThrGluArgTyrMetAspValProGluAsnAsnGlnHisGlyTyrGluAlaGlySerVal 420
QY 1136 GCATCCAGCAGCAAAAGTTCCCTCTCAACCAATCGTTTACTCTCTTACATGGTTTC 1195
Db 421 AlaLeuHisValGluLysLeuProAsnGluProAsnArgLeuLeuIleLeuHisGlyPhe 440
QY 1196 CTGGATGAGAATGTCCATTTTGCACATACCATATATATCTAGTATTTTAGTAGGGCT 1255
Db 441 LeuAspGluAsnValHisPhePheHisThrAsnPheLeuValSerGlnLeuIleArgAla 460
QY 1256 GGAAGCCATATGATTACAGATCTCTCTCAGGAGACACACAGCATAAGAGTTCTCCGAA 1315
Db 461 GlyLysProTyrGlnLeuGlnIleTyrProAsnGluArgHisSerIleLeuArgCysProGlu 480
QY 1316 TCGGAGACATTATGAACTGCATCTTTTGCATCTCTTTCAGACTCTTCAAGAAACCTT 1366
Db 481 SerGlyGluHisTyrGluValThrLeuLeuHisPheLeuGlnGluTyrLeu 497

RESULT 3
US-11-264-096-1593
; Sequence 1593, Application US/11264096
; Publication No. US20060084794A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF546D1
; CURRENT APPLICATION NUMBER: US/11/264,096
; CURRENT FILING DATE: 2005-11-02
; PRIOR APPLICATION NUMBER: 09/833,245
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229,358
```

```
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256, 931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199, 384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 2267
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1593
; LENGTH: 497
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (183)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-11-264-096-1593

Alignment Scores:
Pred. No.: 9,52e-117 Length: 497
Score: 1294.50 Matches: 247
Percent Similarity: 61.2% Conservative: 57
Best Local Similarity: 49.7% Mismatches: 92
Query Match: 43.8% Indels: 101
DB: 11 Gaps: 2

US-10-825-632-6 (1-1669) x US-11-264-096-1593 (1-497)

QY 179 ATCTCTAGATCGCTCCAGACTCGCTCAGATAGTGTGATCTCACTGATATATT 238
Db 1 MetPheLeuAspArgProGlnTrpLeuGlnLeuValLeuLeuProAlaLeuPhe 20

QY 239 ATCCAGTAGAAGATGATTATGGAAGAGGAGAGACTATTGAGTCAGTCTGATCT 298
Db 21 IleProSerThrGluAsnGluGlnArgLeuAlaSerAlaArgAlaValProArgAsn 40

QY 299 GTGAGCCCACTAATATCTATGAAGAAACACAGACATCTGGATAAATATCATCATC 358
Db 41 ValGlnProTyrValValTyrGluGluValThrAsnValTrpIleAsnValHisAspIle 60

QY 359 TTTCATGTTTTCCCAAGTAC--GAAGAGAAATGAGTTATTTTTCCTCTGAA 415
Db 61 PheTyrProPheProGlnSerGluGlyGluAspGluLeuCysPheLeuArgAlaAsnGlu 80

QY 416 TGCAAAACAGTTTCCGCTATTATACAAATACATCTATTAAAGAAAGCAATAT 475
Db 81 CysLysThrGlyPheCysHisLeuTyrLysValThrAlaValLeuLysSerGlnGlyTyr 100

QY 476 AAACGATCCAGTGTGGCTGCTGCTCCAGTGATTTCAAGTGCCTATCAAGAGAG 535
Db 101 AspTrpSerGluProPheSerProGlyGluAspGluPheCysProIleLysGluGlu 120

QY 536 ATACCAATTACAGTGTGAATGGAGTTCTTGGCCGGCATGATCTAAATATCAAGTT 595
Db 121 IleAlaLeuThrSerGlyGluTrpGluValLeuAlaArgHisGlySerLysIleTrpVal 140

QY 596 GATGAAGTCAGAGGCTGTATATTGTAAGGCACAAAGACTCCCTTTAGAGCATCAC 655
Db 141 AsnGluGluThrLysLeuValTyrPheGlnGlyThrLysAspThrProLeuGluHis 160

QY 656 CTGTACATGATGATACGTAATCTCTGAGAGGTGACAAAGCTGACCGTGGGTAC 715
Db 161 LeuTyrValValSerTyrGluAlaGlyGluIleValArgLeuThrProGlyPhe 180

QY 716 TCACATTTCTGTCATCAGTCACTGTGACTTCTTTATAAGTAAGTATAGTAACAG 775
Db 181 SerHis***CysSerMetSerGlnAsnPheAspMetPheValSerHisTyrSerSerVal 200

QY 776 AAGAATCCACACTGTGTGCTCCCTTTTACAGCTATCAAGTCTCAAGATGACCCAACTTC 835
Db 201 SerThrProCysValHisValTyrLysLeuSerGlyProAspAspProLeuHis 220

QY 836 AAAACAAAGGAATTTTGGGCCACCAATTTTGGATTTCAGCAGGTCTCTTCTCTGACTACT 895
Db 836 AAAACAAAGGAATTTTGGGCCACCAATTTTGGATTTCAGCAGGTCTCTTCTCTGACTACT
```

```
Db 221 LysGlnProArgPheTrpAlaSerMetMetGluAlaAlaSerCysProAspTyrVal 240
QY 896 CTCCAGAAATTTCTCTTTTGAAGTACTACTGGATTATACATTGTATGGATGCTCTAC 955
Db 241 ProGluIlePheHisPheHisThrArgSerAspValArgLeuTyrGlyMetIleTyr 260
QY 956 AAGCCTCATGATCTACAGCCTGGAAGAAATATCTACTGTGCTGCTTTCATATATGGTGGT 1015
Db 261 LysProHisAlaLeuGlnProGlyLysHisProThrValLeuPheValTyrGlyGly 280
QY 1016 CCTCAG----- 1021
Db 281 ProGlnValGlnLeuValAsnSerPheLysGlyIleLysTyrLeuArgLeuAsnThr 300
QY 1021 ----- 1021
Db 301 LeuAlaSerLeuGlyTyrAlaValValIleAspGlyArgGlySerCysGlnArgGly 320
QY 1021 ----- 1021
Db 321 LeuArgPheGluGlyAlaLeuLysAsnGlnMetGlyGlnValGluIleGluAspGlnVal 340
QY 1021 ----- 1021
Db 341 GluGlyLeuGlnPheValAlaGluLysTyrGlyPheIleAspLeuSerArgValAlaIle 360
QY 1021 ----- 1021
Db 361 HisGlyTrpSerTyrGlyPheLeuSerLeuMetGlyLeuIleHisLysProGlnVal 380
QY 1022 -----GTTGCTATTGCTGGGGCCCGCTACTCTGTGGATCTTCTATGATACAGGATAC 1075
Db 381 PheLysValAlaIleAlaGlyAlaProValThrValTrpMetAlaTyrAspThrGlyTyr 400
QY 1076 ACGGAACGTTATATGGTTCACCTGACCAAGAACAGAGGCTATTACTTAGGATCTGTG 1135
Db 401 ThrGluArgTyrMetAspValProGluAsnAsnGlnHisGlyTyrGluAlaGlySerVal 420
QY 1136 GCCATCAAGCAGAAAGTTCCTCTGAAACCAATCGTTTACTGTCTTACATGTTTC 1195
Db 421 AlaLeuHisValGluLysLeuProAsnGluProAsnArgLeuLeuIleLeuHisGlyPhe 440
QY 1196 CTGGATGAGAAATGCCATTTTGACATPACCAGTATATTACTAGTGTCTTTAGTAGGGCT 1255
Db 441 LeuAspGluAsnValHisPhePheHisThrAsnPheLeuValSerGlnLeuIleArgAla 460
QY 1256 GGAAGCCCATATGATTTACAGATCTATCCTCAGAGAGACACAGACATAAGATTCCTGAA 1315
Db 461 GlyLysProTyrGlnLeuGlnIleTyrProAsnGluArgHisSerIleArgCysProGlu 480
QY 1316 TCGGAGAGAACATTATGAACATGCACTCTTTGCACTACCTTCAAGAAACCTT 1366
Db 481 SerGlyGluHisTyrGluValThrLeuLeuHisPheLeuGlnGluTyrLeu 497
```

```
RESULT 4
US-11-264-096-1591
; Sequence 1591, Application US/11264096
; Publication No. US20060084794A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PFS46D1
; CURRENT APPLICATION NUMBER: US/11/264,096
; CURRENT FILING DATE: 2005-11-02
; PRIOR APPLICATION NUMBER: 60/933,245
; PRIOR FILING DATE: 2005-04-12
; PRIOR APPLICATION NUMBER: 60/229,359
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 2267
```

```

; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1591
; LENGTH: 349
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (183)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (191)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (334)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (344)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (345)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (348)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; US-11-264-096-1591

Alignment Scores:
Pred. No.: 1,5e-48 Length: 349
Score: 588.50 Matches: 111
Percent Similarity: 72.0% Conservative: 33
Best Local Similarity: 55.5% Mismatches: 55
Query Match: 19.9% Indels: 1
DB: 11 Gaps: 1

US-10-825-632-6 (1-1669) x US-11-264-096-1591 (1-349)
Qy 179 ATCTACTAGATCGCTCCAGACTCGCCTCAGACATAGTGTGTGATCTCACCCTGAATATTT 238
Db 1 MetPheLeuAspArgProGlnGlnTrpLeuGlnLeuValLeuLeuProProAlaLeuPhe 20
Qy 239 ATCCAGTAGAAGATGATTTATGGAAGAGGAGAGACTATTGAGTCAGTCGCTGATTC 298
Db 21 IleProSerThrGluAsnGluGlnGlnArgLeuAlaSerAlaArgAlaValProArgAsn 40
Qy 299 GTGACGCCAATAATATCTATCAAGAACACAGACATCTGGATAAATATCCATGCATC 358
Db 41 ValGlnProTrpValValTrpGluGluValThrAsnValTrpIleAsnValHisAspIle 60
Qy 359 TTTTCATGTTTTCCTCCAAAGTCAC--GAAGAGAAATGAGTTTATTTTTCCTCTGAA 415
Db 61 PheTyProPheProGlnSerGluGlyGluAspGluLeuCysPheLeuArgAlaAsnGlu 80
Qy 416 TDCAAACAGGTTTCCTCATTTATATACAAATTTACATCTATTTTAAAGGAAAGCAATAT 475
Db 81 CysIysThrGlyPheCysHisLeuTyrylsValThrAlaValLeuLysSerGlnGlyTy 100
Qy 476 AAACAGATCCAGTGGTGGCTGCTGCTCCAGTGAATTTCAAGTGAATTCCTATCAAGAGGAG 535
Db 101 AspTrpSerGluProPheSerProGlyGluAspGluPheLysCysProIleLysGluGlu 120
Qy 536 ATAGCAATACAGTGGTGAATGGGAAGTTCTTGGCCGGCATGGATCTAATATCAAGTT 595
Db 121 IleAlaLeuThrSerGlyGluTrpGluValLeuAlaArgHisGlySerLysIleTrpVal 140
```

```

Qy 596 GATGAAGTCAGAAAGCTGCTATATTTTGAAGGCACCAAGACTCCCTTTAGAGCATCAC 655
Db 141 AsnGluGluThrLysLeuValTyPheGlnGlyThrLysAspThrProLeuGluHisHis 160
Qy 656 CTGTACGTAGTCAGTTAGTAAATCCTCAGAGGTGACAAGGCTGACTGACCGTGGCTAC 715
Db 161 LeuTyValValSerTyGluAlaAlaGlyGluIleValArgLeuThrThrProGlyPhe 180
Qy 716 TCACATTCTTGTGTCATCAGTCAGTCAGTCTTCTTTATTAAGTAAGTATAGTAAACAG 775
Db 181 SerHis***CysSerMetSerGlnAsnPhe*****PheValSerHisIleThrAlaGln 200
RESULT 5
US-11-176-951-10
; Sequence 10, Application US/11176951
; Publication No. US20060024313A1
; GENERAL INFORMATION:
; APPLICANT: CHEN, XIN
; TITLE OF INVENTION: AGENTS THAT DISRUPT DIMER FORMATION IN DPP-IV FAMILY OF
; FILE OF INVENTION: PROLYL DIPEPTIDASES
; FILE REFERENCE: 08842.0019
; CURRENT APPLICATION NUMBER: US/11/176,951
; CURRENT FILING DATE: 2005-07-06
; PRIOR APPLICATION NUMBER: 60/586,095
; PRIOR FILING DATE: 2004-07-06
; PRIOR APPLICATION NUMBER: 60/585,952
; PRIOR FILING DATE: 2004-07-06
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 10
; LENGTH: 109
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-11-176-951-10

Alignment Scores:
Pred. No.: 1,39e-47 Length: 109
Score: 577.00 Matches: 109
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 19.5% Indels: 0
DB: 11 Gaps: 0

US-10-825-632-6 (1-1669) x US-11-176-951-10 (1-109)
Qy 1070 GGATACACGAACGTTTATATGGTCACCTGACCAAGATGAACAGGGCTATTACTTAGGA 1129
Db 1 GlyTyThrGluArgTyMetGlyHisProAspGlnAsnGluGlnGlyTyTyLeuGly 20
Qy 1130 TCTGTGCCATGCAACAGCAAAAAGTTCCCTCTGACCAAAATCGTTTACTGCTCTTACAT 1189
Db 21 SerValAlaMetGlnAlaGlyLysPheProSerGluProAsnArgLeuLeuLeuHis 40
Qy 1190 GGTTCCTCGATGAGATCTCCATTTTGCACATACACATATATTACTGAGTTTTTTAGTG 1249
Db 41 GlyPheLeuAspGluAsnValHisPheAlaHisThrSerIleLeuLeuSerPheLeuVal 60
Qy 1250 AGGCTGGAAGGCATATCATTTTACAGATCTATCTCTCAGGAGAGACACAGCATAGAGTT 1309
Db 61 ArgAlaGlyLysProTyAspLeuGlnIleTyProGlnGluArgHisSerIleArgVal 80
Qy 1310 CCTGAATCGGGAGAACATTATGAACATGCTCTTTTGCACACTACCTTCAAGAAAACCTTGA 1369
Db 81 ProGluSerGlyGluHisTyGluLeuHisLeuLeuHisLeuGlnGlnLeuGlnLeuGly 100
Qy 1370 TCACGATTCTCTCTTAAAGTGATA f396
Db 101 SerArgIleAlaAlaLeuLysValIle 109
RESULT 6
US-11-176-951-11
; Sequence 11, Application US/11176951
```

```
; Publication No. US20060024313A1
; GENERAL INFORMATION:
; APPLICANT: CHEN, YUAN, SHOU
; TITLE OF INVENTION: AGENTS THAT DISRUPT DIMER FORMATION IN DPP-IV FAMILY OF
; FILE REFERENCE: PROLYL DIPEPTIDASES
; CURRENT APPLICATION NUMBER: US/11/176,951
; CURRENT FILING DATE: 2005-07-06
; PRIOR APPLICATION NUMBER: 60/586,095
; PRIOR FILING DATE: 2004-07-06
; PRIOR APPLICATION NUMBER: 60/585,952
; PRIOR FILING DATE: 2004-07-06
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 11
; LENGTH: 99
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-176-951-11

Alignment Scores:
Pred. No.: 6,62e-30 Length: 99
Score: 394.00 Matches: 72
Percent Similarity: 82.8% Conservative: 10
Best Local Similarity: 72.7% Mismatches: 17
Query Match: 13.3% Indels: 0
DB: 11 Gaps: 0

US-10-825-632-6 (1-1669) x US-11-176-951-11 (1-99)
QY 1070 GGATACACGAGCGTATATGGTCCACCTGCACAGATGAACAGCGCTATTACTTAGGA 1129
DB 1 GlyTyrThrGluArgTyrMetAspValProGluAsnGlnHisGlyTyrGluAlaGly 20
QY 1130 TCTGTGCGCATCAAGCAGAGAAAGTCCCTCTGAAACAAATCGTTTACTGCTTTACAT 1189
DB 21 SerValAlaLeuHisValGluLysLeuProAsnGluProAsnArgLeuLeuLeuHis 40
QY 1190 GGTTCCTCGATGAGAATGTCCATTTTGCACATACCAAGTATATTACTGAGTTTGTAGT 1249
DB 41 GlyPheLeuAspGluAsnValHisPheHisThrAsnPheLeuValSerGlnLeuLeu 60
QY 1250 AGGCTGGAAAGCCATATGATTTACAGATCTATCTCTCAGGAGAGACACAGCATAGAGTT 1309
DB 61 ArgAlaGlyLeuProTyrGlnLeuGlnIleTyrProAsnGluArgHisSerIleArgCys 80
QY 1310 CTGTAATCGGGAGAACATTATGAATGCATCTTTTGGCACTACCTCTTCAAGAAACCTT 1366
DB 81 ProGluSerGlyGluHisTyrGluValThrLeuLeuHisPheLeuGlnGluTyrLeu 99

RESULT 7
US-11-079-463-6408
; Sequence 6408, Application US/11079463
; Publication No. US20060073161A1
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTERIOIDES FR
; FILE REFERENCE: PATH00-03DIV2
; CURRENT APPLICATION NUMBER: US/11/079,463
; CURRENT FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/128,705
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: US 09/540,209
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 10444
; SEQ ID NO 6408
; LENGTH: 745
; TYPE: PRT
; ORGANISM: B. fragilis
US-11-079-463-6408
```

```
Alignment Scores:
Pred. No.: 2,54e-26 Length: 745
Score: 359.50 Matches: 125
Percent Similarity: 36.4% Conservative: 74
Best Local Similarity: 22.9% Mismatches: 163
Query Match: 12.2% Indels: 185
DB: 11 Gaps: 17

US-10-825-632-6 (1-1669) x US-11-079-463-6408 (1-745)
QY 20 AAAAGTCACATTTTAAAGATGCAGAAATAATGATGTGCTGAAGGAAGGATCATAGATGTC 79
DB 286 LysValThr---ArgGlnValLysLeuProIleAspAlaAspGly- 299
QY 80 ATAGATAAGGAACATAATTCACCTTTTGAGATTCTATTGAAGGAGTTGAATATATTGCC 139
DB 300 -----TyIlePro 302
QY 140 AGAGCTGGATGGACTCCTCGAGGAAAATATGCTTGGTCCATCTACTAGATCGCTCCAG 199
DB 303 ArgIleArgPheThrGlnAspProAsnLysLeuAlaIleMetThrLeuAsnArgHisGln 322
QY 200 ACTCGCTACAGATAGTGTGATCTCAGCTGAATATTATTATCCAGTAGAAGATGATGTT 259
DB 323 AsnArgPheAspMetTyrPheAlaAspPro 332
QY 260 ATGGAAGGCGAGAGACTCATTCAGTCAGTCTGATCTCTGTGAGCCACTAATATTCTAT 319
DB 333 -----ArgSerThrValCysLysLeuAlaLeuArgAspGluSerProTyrTyrIleAsn 350
QY 320 GAAGAAACACAGACATCTGGATATAATATCATGACATCTTTTCATGTTTTCCTCCAAAGT 379
DB 351 Glu-----AsnValPheAspAsnIleGlnPheTyrProGlu--- 362
QY 380 CACGAAGAGAAATTGAGTTTATTTTTCCTCTCTGAATGCAAAACAGGTTTCCTGTCATTTA 439
DB 363 -----TyrPheSerPheValSerAsp---LysSerGlyTyrProHisLeu 376
QY 440 TACAAATATACATCTATTATTAAGGAAAGCAATATAACGATCCAGTGTGGGCTGCCT 499
DB 377 Tyr-----TrpTyrSerMetAsnGlyAsnLeu--- 385
QY 500 GCTCCAAGTGATTCCAAGTGCTCTATCAAGAGAGAGATAGCAATTACAGTGGTGAATGG 559
DB 386 -----IleLysGln-----ValThrSerGlyAsnTyr 394
QY 560 GAAGTTCTTGGCGCGCATGGATCTAATATCCAAGTTGATGAAGTCAGAGGCTGGTATAT 619
DB 395 GluValLysAsnPheIleGlyTyrAsnProAspThrAsnGlu-----PheTyr 410
QY 620 TTGGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTAGTACGTACGTACGTAAAT 679
DB 411 TyrThrSerAsnGluSerProMetArgGlnAlaValTyrLysIleAsp-----Arg 428
QY 680 CCTGGAGAGGTGACAGGCTGACTGACCGTGGCTACTCAGATTTCTTCTGCATCAGTCAG 739
DB 429 LysGlyLysLysMetLysLeuSerAsnGlnProGlyThrAsnSerProIlePheSerSer 448
QY 740 CACTGTGACTCTTTTAAAGTAAAGTATAGTAACAGAGAAATCCACACTGTGTGCTCCTT 799
DB 449 SerMetLysTyrPheMetAsnLysPheThrSerLeuAspThrProMetLeuIleThrLeu 468
QY 800 TACAAGCTATCAAGTCTCTGAAGAGTACCCCACTTGGCAAAACAAAGGAATTTTGGGCCACC 859
DB 469 -----AsnAspAsnThrGlyLysValLeuLysThrLeuValThr 481
QY 860 ATTTTGGATTACAGAGGCTCTCTTCTGACTATATCTCTCCA-----GAAATTTTCTCT 913
DB 482 AsnAspLysLeuLysGlnLysLeuAlaGluTyrAlaIleProGlnLysGluPhePheThr 501
QY 914 TTTGAAAGTACTACTGAGTTTACATTTGATGGGATGCTCTACAGGCTCATGATCTACAG 973
DB 502 PheLysThrThrGluGlyValAspLeuAsnGlyTyrMetMetLysProValAsnPheAsp 521
```



```
Db 435 TyrMetAlaAsnLysGlyTyrIleMetPheThrValAspGlyArgGlySerSerAsnArg 454
Qy 1018 -----
Db 455 GlyLeuAspPheGluAsnValThrPheArgGlnLeuGlyIleGluGluGlyArgAspGln 474
Qy 1018 -----
Db 475 ValLysGlyThrGluPheLeuLysSerLeuProTyrValAspGlyAsnArgIleGlyVal 494
Qy 1018 -----
Db 495 HisGlyTrpSerPheGlyHisMetThrThrLalaLeuLeuLeuArgTyrProGluIle 514
Qy 1019 ---CAGTTGCTATTGCTGGGCCCCAGTCACTCTGTGGATCTTCTTAAGTACAGGATAC 1075
Db 515 PheLysValGlyValAlaGlyGlyProValIleAspTrpGlyTyrTyrGluValMetTyr 534
Qy 1076 ACGAAACGTTATATGGTCAACCTGACCAAGTAAACAGAGGCTATTACTTTAGGATCTGTG 1135
Db 535 GlyGluArgTyrMetAspThrProGlnSerAsnProLysGlyTyrLysGluCysAsnLeu 554
Qy 1136 GCATGCAAGCAAGAAAGTTCCCTCTGAACCAATCGTTACTGCTCTTACATGGTTTC 1195
Db 555 LysAsnLeuAlaGlyAsnLeu-----LysGlyHisLeuMetIleHisAspAsp 571
Qy 1196 CTGGATGAGATGTCCTCAITTTGCACATACCATATATTACTGAGTTTCTTGTAGGAGGCT 1255
Db 572 HisAspAspThrCysValProGlnHisThrLeuSerPheMetLysAlaCysIleAspAla 591
Qy 1256 GGAAGGCCATATGATTTACAGATCTATCTCAGGAGAGACACAGCATAAGAGTTCTCTGAA 1315
Db 592 ArgThrTyrProAspLeuPheIleTyrProCysHisLysHisVal----- 607
Qy 1316 TCGGAGAA-----CATTATGAAGTCACTTTTGGTACCTTTCACCTTCAAGAAACCTT 1366
Db 608 SerGlyArgAspArgValHisLeuHisGluLysIleThrArgTyrPheGluAspTyrLeu 627

RESULT 9
US-11-208-288-4
; Sequence 4: Application US/11208288
; Publication No: US20060051366A1
; GENERAL INFORMATION
; APPLICANT: CHANG, Chiwen
; TITLE OF INVENTION: USE OF SOLUBLE CD26 AS INHIBITOR OF
; FILE OF INVENTION: ANGIOGENESIS AND INFLAMMATION
; FILE REFERENCE: 39533-0001
; CURRENT APPLICATION NUMBER: US/11/208,288
; PRIOR FILING DATE: 2005-08-18
; PRIOR APPLICATION NUMBER: US 60/605,013
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 738
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-208-288-4

Alignment Scores:
Pred. No.: 1,4e-14 Length: 738
Score: 238.00 Matches: 120
Percent Similarity: 33.5% Conservative: 71
Best Local Similarity: 21.1% Mismatches: 193
Query Match: 8.0% Indels: 186
DB: 11 Gaps: 24

US-10-825-632-6 (1-1669) x US-11-208-288-4 (1-738)

Qy 5 GGTACACCAATCTTAAGTCACTTTTAAGATGTCAGAAATTAATGATTGCTGAAGGA 64
Db 232 GlyAlaValAsnProThrValLysPhePheVal-----ValAsnThrAspSerLeuSer 249
```

```
Qy 65 AGGATCATAGATGCTCATAGATAGGAACCTAAATCAACCTTTTGAGATTCTATTGAAGGA 124
Db 250 SerValThrAsnAlaThrSerIleGlnIleThrAlaProAlaSerMetLeuIle---Gly 268
Qy 125 GTTGAATATATATCCAGAGCTGGATGAGCTCTCTGAGGAAAAATATGCTTGGTCCATCCTA 184
Db 269 AspHisTyrLeuCysAspValThrTrpAla----- 278
Qy 185 CTAGATCGCTCCAGACTCGCTCATAGATAGTGTGATCTCACCTGAATATTATTATCCCA 244
Db 279 -----ThrGlnGluArgIleSerLeuGlnTrpLeu----- 288
Qy 245 GTAGAAAGATGATGTTATGAAAGCGAGAGACTCAATTGAGTCAGTGCCTGATTCTGTGACG 304
Db 289 -----ArgArgIleGlnAsnTyr-----SerValMet 297
Qy 305 CCACTAATTATCTATGAAGAAACACAGACATCTGGATAAATATCCATGACATCTTTCAT 364
Db 298 AspIleCysAspTyrAspGluSerSerGlyArgTrp---AsnCys----- 311
Qy 365 GTTTTCCCAAGTACAGAGAGAAATGAGTTTATTTTGGCTCTGAATGCAAAACA 424
Db 312 LeuValAlaArgGlnHisIleGluMetSerThrThrGlyTrpValGlyArgPheArgPro 331
Qy 425 GGTTCCTCGTCAT-----TTATACAAAATTACATCTATTTTAAAG 463
Db 332 SerGluProHisPheThrLeuAspGlyAsnSerPheTyrLysIleLeuSer-----Asn 349
Qy 464 GAAGCAAAATATAACGATCCAGTGGTGGCTGCTCTCCAGTGATTTTCAAGTGTCT 523
Db 350 GluGluGlyTyrArgHisIle-----CysTyrPheGlnIleAsp 362
Qy 524 ATCAAGAGAGATAGCAATACAGTGGTGAATGGGAAGTTCTTGGCCGCGATGGATCT 583
Db 363 LysLysAspCysThrPheIleThrLysGlyThrTrpGluValIleGly----- 378
Qy 584 AATATCCAAGTGTGATGAAGCTGATATATTTTGAAGGCACC-----AAAGACTCC 640
Db 379 ---IleGluAlaLeuThrSerAspTyrLysTyrIleSerAsnGluTyrLysGlyMet 397
Qy 641 CCTTTAGAGCATCACCTGTAGTACGATTACCTAAATCTCGAGAGGTCAGCAAGGCTG 700
Db 398 ProGlyGlyArgAsnLeuTyrLysIleGlnLeuIleAspTyrThrLysValThrCysLeu 417
Qy 701 ACTGACCGTGGCTACTCACAATCTTGTGTCATCAGT---CAGCACGTGTGATCTTTTATA 757
Db 418 -----SerCysGluLeuAsnProGluArgCysGlnTyrTyrSer 430
Qy 758 AGTAAGTATAGTAAACCAAGAAATCCACAC-----TGT-----GTGTCC 796
Db 431 ValSerPheSerLysGluAlaLysTyrTyrGlnLeuArgCysSerGlyProGlyLeuPro 450
Qy 797 CTTTACAAGCTATCAAGTCTGAAGATGACCCAACTTGCAAAACAAGGAATTTGGGCC 856
Db 451 LeuTyrThrLeuHisSerSerValAsnAspLysGlyLeuArgValLeuGluAsp---Asn 469
Qy 857 ACCATTTTGGATTGAGCAGAGGCTCTTCTGCTGATCTACTCTCTCCAGAA-----ATT 907
Db 470 SerAlaLeuAspLysMet-----LeuGlnAsnValGlnMetProSerLysLysLeuAsp 487
Qy 908 TTCTCTTTTGAATACTACTGGATTACATTTGATGGGATGCTCTACAAGCTCATGAT 967
Db 488 PheIleIleLeuAsnGluThrLysPhe---TrpTyrGlnMetIleLeuProHis--- 505
Qy 968 CTACAGCTGGAAAGAAATATCTCTACTGTCTCTTATATATGTTGGTGGTCTCT 1018
Db 506 PheAspLysSerLysTyrProLeuLeuAspValTyrAlaGlyProCysSerGln 525
Qy 1018 ----- 1018
Db 526 LysAlaAspThrValPheArgLeuAsnTrpAlaThrTyrLeuAlaSerThrGluAsnIle 545
```


QY	1018	-----	1018
Db	546	IleValAlaSerPheAspGlyArgGlySerGlyTyrGlnGlyAspLysIleMetHisAla	565
QY	1018	-----	1018
Db	566	IleAsnArgArgLeuGlyThrPheGluValGluAspGlnIleGluAlaAlaArgGlnPhe	585
QY	1018	-----	1018
Db	586	SerLysMetGlyPheValAspAsnLysArgIleAlaIleTrpGlyTrpSerTyrGlyGly	605
QY	1019	-----CAGGTTCCTATTGCTGGG	1036
Db	606	TyrValThrSerMetValLeuGlySerGlySerGlyValPheLysCysGlyIleAlaVal	625
QY	1037	GCCCACTCACTCTGTGGATCTCTATGATACAGATACACGGAACGTTATATGGGT---	1093
Db	626	AlaProValSerArgTrpGluTyrTyrAspSerValTyrThrGluArgTyrMetGlyLeu	645
QY	1094	---CACCTCACCAGATGAACAGGCTATTACTTAGATCTGTGGCCATCAAGCAGAA	1150
Db	646	ProThrProGluAspAsnLeuAspHisTyrArgAsnSerThrValMetSerArgAlaGlu	665
QY	1151	AGTTCCCTCTGAACCAATCGTTTACTGCTCTTACATGTTTCTCGATGAGAATGTC	1210
Db	666	AsnPhe-----LysGlnValGluTyrLeuLeuIleHisGlyThrAlaAspAsnVal	683
QY	1211	CATTTCGACATACAGTATATTAAGTCTTTTGTAGTGGGCTGGAAAGCCATATGAT	1270
Db	684	HisPheGlnGlnSerAlaGlnIleSerLysAlaLeuValAspValGlyValAspPheGln	703
QY	1271	TACAGATCTATCTCAGGAGACACACAGATAGATTCCTGAATCGGGAGACATTAT	1330
Db	704	AlaMetTrpTyrThrAspGluAspHisGlyIleAlaSerSerThrAlaHisGlnHisle	723
QY	1331	GAACGTCACTTTTGCCTACTCTCAAGAA	1360
Db	724	TyrThrHisMetSerHisPheIleLysGln	733
RESULT 10			
US-10-501-035-234			
; Sequence 234, Application US/10501035			
; Publication No. US20060046249A1			
; GENERAL INFORMATION:			
; APPLICANT: Bristol-Myers Squibb Company			
; TITLE OF INVENTION: IDENTIFICATION OF POLYNUCLEOTIDES AND POLYPEPTIDE FOR PREDICTING			
; TITLE OF INVENTION: ACTIVITY OF COMPOUNDS THAT INTERACT WITH PROTEIN TYROSINE KINASE			
; TITLE OF INVENTION: AND/OR PROTEIN TYROSINE KINASE PATHWAYS			
; FILE REFERENCE: D0185 PCT			
; CURRENT APPLICATION NUMBER: US/10/501,035			
; CURRENT FILING DATE: 2004-07-09			
; PRIOR APPLICATION NUMBER: US 60/350,061			
; PRIOR FILING DATE: 2002-01-18			
; NUMBER OF SEQ ID NOS: 795			
; SOFTWARE: PatentIn version 3.2			
; SEQ ID NO 234			
; LENGTH: 766			
; TYPE: PRT			
; ORGANISM: Homo sapiens			
US-10-501-035-234			
Alignment Scores:			
Pred. No.:	1,42e-14	Length:	766
Score:	238.00	Matches:	120
Percent Similarity:	33.5%	Conservative:	71
Best Local Similarity:	21.1%	Mismatches:	193
Query Match:	8.0%	Indels:	186
DB:	9	Gaps:	24
US-10-825-632-6 (1-1669) x US-10-501-035-234 (1-766)			
QY	5	GGTACAGCAATCCTTAAGTCATCTTTTAAGATGTCAGAAATAATGATGATGTCGAAGGA	64

Db	260	GlyAlaValAsnProThrValLysPhePheVal-----ValAsnThrAspSerLeuSer	277
QY	65	AGGATCATAGATGTCATAGATAAGGAACCTAAATTCACCTTTTGTAGATCTTATTTGAAGGA	124
Db	278	SerValThrAsnAlaThrSerIleGlnIleThrAlaProAlaSerMetLeuIle---Gly	296
QY	125	GTTCAATATATTTGCCAGAGCTGGATGAGTCTCTGAGGGAAATATATGTTGGTCCATCCTA	184
Db	297	AspHisTyrLeuCysAspValThrTrpAla-----	306
QY	185	CTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAATTTATTTATCCCA	244
Db	307	-----ThrGlnGluArgIleSerLeuGlnTrpLeu-----	316
QY	245	GTAAGAATGATGTTTGGAAAGGACAGAGACTCAATGAGTCAGTCGCTGATTTCTGTGACG	304
Db	317	-----ArgGlnIleGlnAsnTyr-----SerValMet	325
QY	305	CCACTAATTATCTATGAAGAAACACAGACATCTGGATAAATATCCATGACATCTTTTCAT	364
Db	326	AspIleCysAspTyrAspGluSerSerGlyArgTrp---AsnCys-----	339
QY	365	GTTTTCCTCCCAAGTACAGAGAGAAATTTAGTATTTATTTTGGCTCTGAATGCAAAACA	424
Db	340	LeuValAlaArgGlnHisIleGluMetSerThrThrGlyTrpValGlyArgPheArgPro	359
QY	425	GTTTTCCTCCAT-----TTATACAAATTTACATCTATTTTAAAG	463
Db	360	SerGluProHisPheThrLeuAspGlyAsnSerPheTyrLysIleIleSer-----Asn	377
QY	464	GAAGCAATATATAACATCCAGTGGTGGCTGCTCCAGTATTTCAAGTGTCTCT	523
Db	378	GluGluGlyTyrArgHisIle-----CysTyrPheGlnIleAsp	390
QY	524	ATCAAGAGGAGATAGCAATACCAAGTGGTGAATCTTCCGCGCATGATCT	583
Db	391	LysLysAspCysThrPheIleThrLysGlyThrTrpGluValIleGly-----	406
QY	584	AATATCCAAAGTTGATGAAGTCAGAGGCTGTATATTTTGAAGGCACC---AAAGACTCC	640
Db	407	-----IleGluAlaLeuThrSerAspTyrLeuTyrIleSerAsnGluTyrLysGlyMet	425
QY	641	CCTTTAGAGCATCACCTGTAGTACCTAGTTACCTAAATCTCGAGAGGTGACAAAGGCTG	700
Db	426	ProGlyGlyArgAsnLeuTyrLysIleGlnLeuIleAspTyrThrLysValThrCysLeu	445
QY	701	ACTGACCGTGGTACTCACTTTTGTGTCATAGT---CAGCACTGTGATCTTTTATA	757
Db	446	-----SerCysGluLeuAsnProGluArgCysGlnTyrTyrSer	458
QY	758	AGTAAGTATAGTAACAGAGATCCACAC-----TGT-----GTGTCC	796
Db	459	ValSerPheSerLysGluAlaLysTyrTyrGlnLeuArgCysSerGlyProGlyLeuPro	478
QY	797	CTTTACAAGCTATCAAGTCCCTGAAGATGACCCCAACTTGCACAAACAAAGAAATTTGGGCC	856
Db	479	LeuTyrThrLeuHisSerValAsnAspLysGlyLeuArgValLeuGluAsp---Asn	497
QY	857	ACCATTTTGGATTACAGAGGTCCTCTTCTGACTATATCTCTCCAGAA-----ATT	907
Db	498	SerAlaLeuAspLysMet-----LeuGlnAsnValGlnMetProSerLysLysLeuAsp	515
QY	908	TTCTCTTTTGAAGTACTACTGATTTGATGGATGCTCTACAAAGCTCATGAT	967
Db	516	PheIleIleLeuAsnGluThrLysPhe---TrpTyrGlnMetIleLeuProProHis---	533
QY	968	CTACAGCCTGGAAGAAATATCTACTGTGCTGTTTCATATATGTTGTCCT	1018
Db	534	PheAspLysSerLysLysTyrProLeuLeuLeuAspValTyrAlaGlyProCysGln	553
QY	1018	-----	1018

```
Db 554 LysAlaAspThrValPheArgLeuAsnTrpAlaThrTyrLeuAlaSerThrGluAsnIle 573
Qy 1018 -----
Db 574 IleValAlaSerPheAspGlyArgGlySerGlyTyrGlnGlyAspLysIleMetHisAla 593
Qy 1018 -----
Db 594 IleAsnArgArgLeuGlyThrPheGluValGluAspGlnIleGluAlaAlaArgGlnPhe 613
Qy 1018 -----
Db 614 SerLysMetGlyPheValAspAsnLysArgIleAlaIleTrpGlyTrpSerTyrGlyGly 633
Qy 1019 -----
Db 634 TyrValThrSerMetValLeuGlySerGlySerGlyValPheLysCysGlyIleAlaVal 653
Qy 1037 GCCCAGTCACCTCTGTCATCTTCTATGATACAGGATACACGGAAGTTATATGGT--- 1093
Db 654 AlaProValSerArgTrpGluTyrTrpAspValTyrThrGluArgTyrMetGlyLeu 673
Qy 1094 ---CACCTGACCAAGATGAACAGGCTATTACTTAGGATCTGTGCCATGCCAAGCAGAA 1150
Db 674 ProThrProGluAspAsnLeuAspHisTyrArgAsnSerThrValMetSerArgAlaGlu 693
Qy 1151 AAGTTCCTCTGAACCAATCGTTTACTCTTACATGGTTTCTCGATCAGAAATGTC 1210
Db 694 AsnPhe-----LysGlnValGluTyrLeuLeuIleHisGlyThrAlaAspAspAsnVal 711
Qy 1211 CATTTGGACATACACGATATATTACTGAGTTTTAGTGAGGCTGGAAAGCCATATGAT 1270
Db 712 HisPheGlnGlnSerAlaGlnIleSerLysAlaLeuValAspValGlyValAspPheGln 731
Qy 1271 TTACAGATCTATCTCCAGGAGACACAGCATAAAGATTCCTGAATCGGAGAACATTAT 1330
Db 732 AlaMetTrpTyrThrAspGluAspHisGlyIleAlaSerSerThrAlaHisGlnHisIle 751
Qy 1331 GAACGTGATCTTTGGCACTACCTCAAGAA 1360
Db 752 TyrThrHisMetSerHisPheIleLysGln 761

RESULT 11
US-11-208-288-2
; Sequence 2, Application US/11208288
; Publication No. US20060051366A1
; GENERAL INFORMATION:
; APPLICANT: CHANG, Chiwen
; TITLE OF INVENTION: USE OF SOLUBLE CD26 AS INHIBITOR OF
; FILE REFERENCE: ANGIOGENESIS AND INFLAMMATION
; CURRENT APPLICATION NUMBER: US/11/208,288
; PRIOR FILING DATE: 2005-08-18
; PRIOR APPLICATION NUMBER: US 60/605,013
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows version 4.0
; SEQ ID NO 2
; LENGTH: 766
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-208-288-2

Alignment Scores:
Pred. No.: 1,42e-14 Length: 766
Score: 238.00 Matches: 120
Percent Similarity: 33.5% Conservative: 71
Best Local Similarity: 21.1% Mismatches: 193
Query Match: 8.0% Indels: 186
DB: 11 Gaps: 24

US-10-825-632-6 (1-1669) x US-11-208-288-2 (1-766)
```

```
Qy 5 GGTACAGCAAACTCTAAAGTCACCTTTTAAGATGTGAGAAATAATGATTGATGCTGAAGGA 64
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
260 GlyAlaValAsnProThrValLysPheVal-----ValAsnThrAspSerLeuSer 277
Qy 65 AGGATCATAGATGTCTATAGTAAGAACTAAATCAACCTTTTGAGATTCTTATTTGAAGGA 124
Db ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: |||||
278 SerValThrAsnAlaThrSerIleGlnIleThrAlaProAlaSerMetLeuIle---Gly 296
Qy 125 GTTGAATATATTGCCAGAGCTGGAGTGCCTCTGAGGGAATAATGCTTGGTCCATCCTA 184
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
297 AspHisTyrLeuCysAspValThrTrpAla----- 306
Qy 185 CTAGATCGCTCCACAGCTCGCTACAGATAGTGTGTGATCTCACCTGAATTATTATCCCA 244
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
307 -----ThrGlnGluArgIleSerLeuGlnTrpLeu----- 316
Qy 245 GTAGAAGATGATGTTATGGAAGGCAGAGACTATTGAGTCAGTCGCTGATCTTGTGACG 304
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
317 -----ArgArgIleGlnAsnTyr-----SerValMet 325
Qy 305 CCACATAATTATCTATGAAGAACACACACATCTGGATAAATATCCATGATCTTTTCAT 364
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
326 AspIleCysAspTyrAspGluSerSerGlyArgTrp---AsnCys----- 339
Qy 365 GTTTTTCCTCCAAAGTCACGAAGAGAAATTGAGTTTATTTTGCCTCTGAAATGCAAAACA 424
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
340 LeuValAlaArgGlnHisIleGluMetSerThrThrGlyTrpValGlyArgPheArgPro 359
Qy 425 GGTTCCTCCGTCAT-----TTATACAAAATTACATCTATTTTAAAG 463
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
360 SerGluProHisPheThrLeuAspGlyAsnSerPheTyrLysIleIleSer-----Asn 377
Qy 464 GAAGCAAAATATAAACGATCCAGTGTGGCTGCTCTCAAGTGATTTCAAGTGTCTCT 523
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
378 GluGluGlyTyrArgHisIle-----CystTyrPheGlnIleAsp 390
Qy 524 ATCAAGAGGAGATAGCAATTTACAGTGTGAATGGAAAGTTCTTGGCGGCATCGATCT 583
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
391 LysLysAspCysThrPheIleThrLysGlyThrTrpGluValIleGly----- 406
Qy 584 AATATCCAAGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGGCACC---AAAGACTCC 640
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
407 ---IleGluAlaLeuThrSerAspTyrLeuTyrTyrIleSerAsnGluTyrLysGlyMet 425
Qy 641 CTTTATAGAGTACCTGTACGTAGTACGTACGTAAATCTCGAGAGAGTGCACAGGCTG 700
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
426 ProGlyGlyArgAsnLeuTyrLysIleGlnLeuIleAspTyrThrLysValThrCysLeu 445
Qy 701 ACTGACCGTGGCTACTCACATTTCTGTCATCACT---CAGCACTGTGACTTCTTTATA 757
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
446 -----SerCysGluLeuAsnProGluArgCysGlnTyrTrpSer 458
Qy 758 AGTAAGTATAGTAACCAAGAAATCCACAC-----TGT-----GTGTCC 796
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
459 ValSerPheSerLysGluAlaLysTyrTyrGlnLeuArgCysSerGlyProGlyLeuPro 478
Qy 797 CTTTCAAGCTATCAGTCTCGAAGATGACCACTTGCAAAACAAGAAATTTTGGGCC 856
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
479 LeuTyrThrLeuHisSerSerValAsnAspLysGlyLeuArgValLeuGluAsp---Asn 497
Qy 857 ACCATTTTGGATTACAGAGCTCTCTCTCACTATATCTCTCCAGAA-----ATT 907
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
498 SerAlaLeuAspLysMet-----LeuGlnAsnValGlnMetProSerLysLysLeuAsp 515
Qy 908 TTCTCTTTTGAAGTACTACTGGATTATCTATGATGGATGCTCTCAACAGCCTCATGAT 967
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
516 PheIleIleLeuAsnGluThrLysPhe---TrpTyrGlnMetIleLeuProHis--- 533
Qy 968 CTACAGCCTGGAAAGAAATATCTCTGCTGTGTTCATATATATGTTGCTCTCT----- 1018
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
534 PheAspLysSerLysLysTyrProLeuLeuAspValTyrAlaGlyProCysSerGln 553
Qy 1018 ----- 1018
```

Db 554 lysAlaAspThrValPheArgLeuAsnTrpAlaThrTyrLeuAlaSerThrGluAsnIle 573
QY 1018 -----
Db 574 lleValAlaSerPheAspGlyArgGlySerGlyTyrGlnGlyAspLysIleMetHisAla 593
QY 1018 -----
Db 594 lleAsnArgArgLeuGlyThrPheGluValGluAspGlnIleGluAlaAlaArgGlnPhe 613
QY 1018 -----
Db 614 SerLysMetGlyPheValAspAsnLysArgIleAlaIleTrpGlyTrpSerTyrGlyGly 633
QY 1019 -----CAGTTTCTATTCTGGG 1036
Db 634 TyrValThrSerMetValLeuGlySerGlySerGlyValPheLysCysGlyIleAlaVal 653
QY 1037 GCCCAGTCACCTGTGGATCTTCTATGATACAGGATACACGGAAGCTTATATGGT--- 1093
Db 654 AlaProValSerArgTrpGluTyrTyrAspSerValTyrThrGluArgTyrMetGlyLeu 673
QY 1094 ---CACCTTGCACAGAAAGACGGCTATTCTTAGGATCTGTGGCCATGCAAGCAGAA 1150
Db 674 ProThrProGluAspAsnLeuAspHisTyrArgAsnSerThrValMetSerArgAlaGlu 693
QY 1151 AAGTTTCCCTCTGACCAAAATCGTTTACTGCTCTTACATGGTTTCCTGGATGAGAAATGTC 1210
Db 694 AsnPhe-----LysGlnValGluTyrLeuLeuIleHisGlyThrAlaAspAsnVal 711
QY 1211 CATTGTGCATACACAGTATATTAAGTATTTTACTGAGGCTGGAAGCCATATGAT 1270
Db 712 HisPheGlnGlnSerAlaGlnIleSerLysAlaLeuValAspValGlyValAspPheGln 731
QY 1271 TTACAGATCTATCCTCAGCAGACACACAGCATAGAGTCTCCTGAATCGGAGAACATTAT 1330
Db 732 AlaMetTrpTyrThrAspGluAspHisGlyIleAlaSerSerThrAlaHisGlnHisIle 751
QY 1331 GAACGTGCATCTTTTGCACCTACCTTCAAGAA 1360
Db 752 TyrThrHisMetSerHisPheIleLysGln 761

RESULT 12
US-11-116-939-13
; Sequence 13, Application US/11116939
; Publication No. US2005026595A1
; GENERAL INFORMATION:
; APPLICANT: Stephen Tomlinson
; TITLE OF INVENTION: TISSUE TARGETED COMPLEMENT MODULATORS
; FILE REFERENCE: 19113.011502
; CURRENT APPLICATION NUMBER: US/11/116,939
; CURRENT FILING DATE: 2005-04-28
; PRIOR APPLICATION NUMBER: 60/565,907
; PRIOR FILING DATE: 2004-04-28
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 762
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence; note=synthetic
; OTHER INFORMATION: construct
US-11-116-939-13

Alignment Scores:
Pred. No.: 3,45e-14 Length: 762
Score: 234.00 Matches: 120
Percent Similarity: 33.1% Conservative: 70
Best Local Similarity: 20.9% Mismatches: 190
Query Match: 7.9% Indels: 194

DB: 11 Gap: 25
US-10-825-632-6 (1-1669) x US-11-116-939-13 (1-762)
QY 5 GGTACAGCAATCTTAAGTCACCTTTTAAGATGTCAGAAATAATGATGATGCTGAAGCA 64
Db 256 GlyAlaValAsnProThrValLysPhePheVal-----ValAsnThrAspSerLeuSer 273
QY 65 AGGATCATAGATGTCATAGATAAGGAACATAATTCAACCTTTTTCAGATTTCTATTGAAGCA 124
Db 274 SerValThrAsnAlaThrSerIleGlnIleThrAlaProAlaSerMetLeuIle---Gly 292
QY 125 GTTGAATATATTCGACAGCTGGATGACACTCCTCAGGAGAAATATGCTGGTCCATCCTA 184
Db 293 AspHisTyrLeuCysAspValThrTrpAla----- 302
QY 185 CTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAAATTATTTATCCCA 244
Db 303 -----ThrGlnGluArgIleSerLeuGlnTrpLeu----- 312
QY 245 GTAGAAGATGATGTTATGAAAGGCAGAGACTCATTCAGTCAGTGCTGATTCCTGACG 304
Db 313 -----ArgArgIleGlnAsnTyr-----SerValMet 321
QY 305 CCACATAATTATCTATGAAGAAACAACACACATCTGGATAAATATCCATGACATCTTTCAT 364
Db 322 AspIleCysAspTyrAspGluSerSerGlyArgTrp---AsnCys----- 335
QY 365 GTTTTTCCTCCCAAGTCACGAAGAGAAATTCAGTTTATTTTGCCTCTGAATCAAAACA 424
Db 336 LeuValAlaArgGlnHisIleGluMetSerThrThrGlyTrpValGlyArgPheArgPro 355
QY 425 GGTTCCTCGTCAT-----TTATACAAATTTACATCTATTTTAAAG 463
Db 356 SerGluProHisPheThrLeuAspGlyAsnSerPheTyrLysIleIleSer-----Asn 373
QY 464 GAAAGCAAAATATAAACAGTCCAGTGGTGGCTGCTGCTCCCAAGTATTTCAAGTGTCT 523
Db 374 GluGluGlyTyrArgHisIle-----CystTyrPheGlnIleAsp 386
QY 524 ATCAAGAGAGATAGCAATACCAATACCAAGTGTGAATGGGAAGTTCTTGGCCGCGATGATCT 583
Db 387 LysLysAspCysThrPheIleThrLysGlyThrTrpGluValIleGly----- 402
QY 584 AATATCCAAGTTGATGAAGTCAGAAAGCTGGTATATATTTTGAAGGCACC-----AAAGACTCC 640
Db 403 ---IleGluAlaLeuThrSerAspTyrLeuTyrTyrIleSerAsnGluTyrLysGlyMet 421
QY 641 CCTTTAGAGCATCACCTGTCAGTACGTACGTACGTAAATCCTCGAGAGGTGACAGGCTG 700
Db 422 ProGlyGlyArgAsnLeuTyrLysIle-----GlnLeu 432
QY 701 ACTGACGGTGGCTACTCACATTTCTGCTGCATCAGT-----CAGCACTGT 745
Db 433 SerAsp-----TyrThrLysValThrCysLeuSerCysGluLeuAsnProGluArgCys 450
QY 746 GACTTCTTTATAGTAAGTATAGTAAACCAAGAAATCCACAC-----TGT----- 790
Db 451 GlnTyrTyrSerValSerPheSerLysGluAlaLysTyrTyrGlnLeuArgCysSerGly 470
QY 791 -----GTGTCTCTTTTCAAGCTATCAAGTCTCGTAAGATGATGCCAACCTTGCAAAACAAG 844
Db 471 ProGlyLeuProLeuTyrThrLeuHisSerSerValAsnAspLysGlyLeuValLeu 490
QY 845 GAATTTGGCCACCATTTTGGATTCAGCAGGTCTCTCTCTGATATATCTCTCCAGAA 904
Db 491 GluAsp---AsnSerAlaLeuAspLysMet-----LeuGlnAsnValGlnMetProSer 507
QY 905 -----ATTTTCTCTTTTGAAGTACTACTGGAATTTACATTTGATGGATGCTCTTAC 955
Db 508 LysLysLeuAspPheIleLeuAsnGluThrLysPhe---TrpTyrGlnMetIleLeu 526
QY 956 AAGCCTCATGATCTACAGCCTGGAAGAAATATCTCTACTGTGCTGCTTTCATATATGGTGT 1015

531	ProProHis---	PheAspLysSerLysLysTyrProLeuLeuLeuAspValTyrAlaGly	549
1016	CCT---		1018
550	ProCysSerGlnLysAlaAspThrValPheArgLeuAsnTrpAlaThrTyrLeuAlaSer	569	
1018	---		1018
570	ThrGluAsnIlelleValAlaSerPheAspGlyArgGlySerGlyTyrGlnGlyAspLys	589	
1018	---		1018
590	IleMetHisAlaIleAsnArgArgLeuGlyThrPheGluValGluAspGlnIleGluAla	609	
1018	---		1018
610	AlaArgGlnPheSerLysMetGlyPheValAspAsnLysArgIleAlaIleTrpGlyTrp	629	
1019	---		1024
630	SerTyrGlyGlyTyrValThrSerMetValLeuGlySerGlySerGlyValPheLysCys	649	
1025	GCTATTGCTGGGGCCAGTCACTCTCTGATCTTATGATACAGGATACAGCGAAGT	1084	
650	GlyIleAlaValAlaProValSerArgTrpGlnTyrTyrAspSerValTyrThrGluArg	669	
1085	TATATGGGT-----		1138
670	TyrMetGlyLeuProThrProGluAspAsnLeuAspHisTyrArgAsnSerThrValMet	689	
1139	ATGCAAGCAAGAAAGTCCCTCTGTAACCAAAATCGTTTACTGCTCTTATCATGGTTTCCTG	1198	
690	SerArgAlaGluAsnPhe-----		707
1199	GATGAGAATGTCATTTTGCACATACAGATATATTACTGAGTTTTTATGAGGGCTGGA	1258	
708	AdpAspAsnValHisPheGlnGlnSerAlaGlnIleSerLysAlaLeuValAspValGly	727	
1259	AAGCCATGATTATTCAGATCTTCTCCTCAGAGAGACACAGCATAGAGATTCCTGAATCG	1318	
728	valAspPheGlnAlaMetTrpTyrThrAspGluAspHisGlyIleAlaSerSerThrAla	747	
1319	GDAGAAACATTATGAATCGATCTTTTGGCATCTTACCTTCAAGAA	1360	
748	HisGlnHisIleTyrThrHisMetSerHisPheIleLysGln	761	

```

RESULT 14
US-11-208-288-6
; Sequence 6, Application US/11209288
; Publication NO. US20060051366A1
; GENERAL INFORMATION:
; APPLICANT: CHANG, Chiwen
; TITLE OF INVENTION: USE OF SOLUBLE CD26 AS INHIBITOR OF
; TITLE OF INVENTION: ANGIOGENESIS AND INFLAMMATION
; FILE REFERENCE: 39533-0001
; CURRENT APPLICATION NUMBER: US/11/208,288
; CURRENT FILING DATE: 2005-08-18
; PRIOR APPLICATION NUMBER: US 60/605,013
; PRIOR FILING DATE: 2004-08-26
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 760
; TYPE: PRT
; ORGANISM: Mus musculus
US-11-208-288-6

Alignment Scores:
Pred. No.: 1,05e-13 Length: 760
Score: 229.00 Matches: 98
Percent Similarity: 33.5% Conservative: 47
Best Local Similarity: 22.6% Mismatches: 120

```

Query Match:	7.7%	Indels:	168
DB:	11	Gaps:	19
US-10-825-632-6 (1-1669) x US-11-208-288-6 (1-760)			
QY	AAAACAGGTTTCGGTCAATTTATAAAATATACATCTATTATAAAGGAACAATATAAA		478
Db	LysAspGlyTyrllysHisIleCysHisPhe-----		381
QY	CGATCCAGTGGTGCGCTCGCTGCCAAGTAGATTCAAG--TGTCCTATCAAAGAGGAG		535
Db	ProLysAspLysLysAspCysThrPhe-----		390
QY	ATACCAATTACCAGTGGTGAATCGGAAGTCTCTGGCCGCATGGATCTAATCATCAAGTT		595
Db	IleThrLysGlyAlatrPGluValIle-----SerIleGluAla		403
QY	GATCAAGTCAGAAGCTGGGTATATTTTGAAGGCACC---AAAGACTCCCCTTTAGAGCAT		652
Db	LeuThrSerAspTyrlLeutyTyrlIleSerAsnGlnTyrlLysGluMetProGlyGlyArg		423
QY	CACCTGACGTGATCAGTTACGTAAATCTCGGAGAGGTGACAAGGCTGACTGACCTGGC		712
Db	AsnLeutyrlLysIle-----GlnLeuthrAsp-----		432
QY	TACTCACATTTCTGCTGCATCAGT-----CAGCACTGTGACTTCTTTTATA		757
Db	HisThrAsnVallLysCysLeuSerCysAspLeuAsnProGluArgCysGlnTyrlTyrlAla		452
QY	AGTAAGTATAGTAAACACAGAAGATCCACAC-----TGT-----GTGTCC		796
Db	ValSerPheSerLysGluAlaLysTyrlTyrlGlnLeuGlyCysTrpGlyProGlyLeuPro		472
QY	CTTTAACAGCTATCAAGTCTCGAAGATGACCCCAACTTCGAAACAAGNAITTTGGGCC		856
Db	LeutyThrLeuHisArgSerThrAspHisLysGluLeuArgValLeuGlu-----		489
QY	ACCATTTTGGATTCAGCAGGT-----CCCTCTCTGACTATATCTCTCCAGAA---		904
Db	AspAsnSerAlaLeuAspArgMetLeuGlnAspValGlnMetProSerLys		506
QY	ATTTTCTCTTTTGAAGTACTGGAATTTACATTTGATGGATGCTCTCAAG		958
Db	LysLeuAspPheIleValLeuAsnGluThrArgPhe--TrpTyrlGlnMetIleLeuPro		525
QY	CCTCATGATCTACAGCTCGGAAGAAATATCTCTACTGCTGTTCATATATGTTGCTCT		1018
Db	PheAspLysSerLysLysTyrlProLeuLeuLeuAspValTyrlAlaGlyPro		544
QY	-----		1018
Db	CysSerGlnLysAlaAspAlaSerPheArgLeuAsnTrpAlaThrTyrlLeuAlaSerThr		564
QY	-----		1018
Db	GluAsnIleIleValAlaSerPheAspGlyArgGlySerGlyTyrlGlnGlyAspLysIle		584
QY	-----		1018
Db	MetHisAlaIleAsnArgArgLeuGlyThrLeuGluValGluAspGlnIleGluAlaAla		604
QY	-----CAGGTTGCTATTGCTGGG-----		1036
Db	ArgGlnPheValLysMetCylPheValAspSerLysArgValAlaIleTrpGlyTrpSer		624
QY	-----		1036
Db	TyrGlyGlyTyrlValThrSerMetValLeuGlySerGlySerGlyValPheLysCysGly		644
QY	-----GCCCACTCACTCTGGATCTTCTATGATACAGGATACACGGAACGTTAT		1087
Db	IleAlaValAlaProValAspArgTrpGlnTyrlAspSerValTyrlThrGluArgTyr		664

```
QY 1088 ATGGGT-----CACCTGACCAAGATGAACAGGCTATTACTTTAGGATCTGTGGCATG 1141
D 665 MetGlyLeuProIleProGluAspAenLeuAepHisTyrArgAenSerThrValMetSer 684
QY 1142 CAAGCAGAAAGTTCCTCTCTCAACAAATCGTTTACCTGCTCTTACATGGTTTCTCGGAT 1201
D 685 ArgAlaGluHisPhe-----LysGlnValGluTyrLeuLeuIleHisGlyThrAlaAep 702
QY 1202 GAGAAATGTCATTTGACATACCACTATATTTAGTGTCTTTTGTAGTGAGGGCTGGAAG 1261
D 703 AspAenValHisPheGlnSerAlaGlnIleSerLysAlaLeuValAepAlaGlyVal 722
QY 1262 CCATATGATTACAGATCTATCTCTCAGGAGACACACAGCATAAGAGTTCTCGAATCGGA 1321
D 723 AppPheGlnAlaMetTrpTyrThrAspGluAepHisGlyIleAlaSerSerThrAlaHis 742
QY 1322 GACATATGAACCTGATCTTTTGACCTACCTCAAGAA 1360
D 743 GlnHisIleTyrSerHisMetSerHisPheLeuGlnGln 755

RESULT 15
US-11-186-284-55
; Sequence 55, Application US/11186284
; Publication No. US20050266493A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Berger, Allison
; APPLICANT: Guillemette, Tracy L.
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Thibodeau, Stephen N.
; APPLICANT: Burgart, Lawrence J.
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF COLON CANCER
; FILE REFERENCE: MP01-029P2RM
; CURRENT APPLICATION NUMBER: US/11/186,284
; CURRENT FILING DATE: 2005-07-21
; PRIOR APPLICATION NUMBER: US/10/301,822
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 60/339,971
; PRIOR FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: US 60/361,978
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/381,988
; PRIOR FILING DATE: 2002-05-20
; NUMBER OF SEQ ID NOS: 228
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 55
; LENGTH: 760
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-11-186-284-55

Alignment Scores:
Pred. No.: 2,95e-12 Length: 760
Score: 214.00 Matches: 87
Percent Similarity: 32.5% Conservative: 53
Best Local Similarity: 20.2% Mismatches: 127
Query Match: 7.2% Indels: 164
DB: 11 Gaps: 13

US-10-825-632-6 (1-1669) x US-11-186-284-55 (1-760)

QY 419 AAAACAGTTTCGGTCAATTTATACAAATATACATCTATTTTAAGGAAGCAATATAAA 478
D 371 LysAspGlyTyrLysHisIleHisTyrIle----- 380
QY 479 CGATCCAGTGGTGGCTGCTCCCAAGTGATTTCAGTGCTCTATCAAGAGGAGATA 538
D 381 -----LysAspThrValGluAsnAlaIle 388
```

```
QY 539 GCAATTACCAAGTGGTGAAGTCTTGTGCCCGCATGGATCTAATATATCAAGTTGAT 598
D 389 GlnIleThrSerGlyLysTrpGluAlaIle-----AsnIle----- 400
QY 599 GAAGTCAGAAGGCTGTGTATATTTGAAGGACCAAGACTCCCTTTTAGACATCACCTG 658
D 401 -----PheArgValThrGlnAepSerLeuPheTyrSerSerAsn 413
QY 659 TAGCTAGTCAGTTPACGTAAATCTCGAGAGGTGACAGGCTGACTGACCGTGGCTACTCA 718
D 414 GluPheGluGluTyrProGlyArgArgAsnIleTyrArgIleSerIleGlySerTrpPro 433
QY 719 CATTTCT---TGCTGCATCAGTCAGCAC-----TGTCATCTTTTATAAGT 760
D 434 ProSerLysLysCysValThrCysHisLeuArgLysGluArgCysGlnTyrTyrThrAla 453
QY 761 AAGTATAGTAACACAGAGAATCCACACTGTGTGCTCCCTTTAC-----AAGCTA 808
D 454 SerPheSerAspTyrAlaLysTyrTyrAlaLeuValCysTyrGlyProGlyIleProIle 473
QY 809 TCAAGTCCTGAAGATGACCAACTTGCACAAACAAAGGAATTTGGGCCACCATTTGGAT 868
D 474 SerThrLeuHisAspGlyArgThrAspGlnGlu-----IleLysIleLeuGlu 489
QY 869 TCAGCAGGTCCTCTTCCTGAC-----TATACTCTCCAGAAATTTCTCT 913
D 490 GluAenLysGluLeuGluAsnAlaLeuLysAsnIleGlnLeuProLysGluGluLys 509
QY 914 TTTGAAAGTACTACTGGATTTCATTTGATGGGATGCTCTACAGCCTCATGATCTACAG 973
D 510 LysLeuGluValAspGluIleThrLeuTrpTyrLysMetIleLeuProGlnPheAsp 529
QY 974 CTGGAAAGAAATATCTCTACTGCTGCTTCATATATGCTGCTCT----- 1018
D 530 ArgSerLysLysTyrProLeuLeuIleGlnValTyrGlyProCysSerGlnSerVal 549
QY 1018 ----- 1018
D 550 ArgSerValPheAlaValAsnTrpIleSerTyrLeuAlaSerLysGluGlyMetValIle 569
QY 1018 ----- 1018
D 570 AlaLeuValAspGlyArgGlyThrAlaPheGlnGlyAspLysLeuLeuTyrAlaValTyr 589
QY 1018 ----- 1018
D 590 ArgLysLeuGlyValTyrGluValGluAspGlnIleThrAlaValArgLysPheIleGlu 609
QY 1018 ----- 1018
D 610 MetGlyPheIleAspGluLysArgIleAlaIleTrpGlyTrpSerTyrGlyGlyTyrVal 629
QY 1019 -----CAGTTGCTATTGCTGGGCCCA 1042
D 630 SerSerLeuAlaLeuAlaSerGlyThrGlyLeuPheLysCysGlyIleAlaValAlaPro 649
QY 1043 GTCACCTGTGGATCTTCTATGATACAGGATACAGCAAGCTATATNGGTCACCT--- 1099
D 650 ValSerSerTrpGluTyrTyrAlaSerValTyrThrGluArgPheMetGlyLeuProThr 669
QY 1100 ---GACCAGATGAACAGGCTTATTACTTAGGATCTGTGCCCATGCAACAGAAAGTTC 1156
D 670 LysAspAepAenLeuGluHisTyrLysAenSerThrValMetAlaArgAlaGluTyrPhe 689
QY 1157 CCTCTGACCAATCTTTACTGCTCTTACATPGTTTCTCGATGAGAAATCTCATTTT 1216
D 690 ArgAenValAsp-----TyrLeuLeuIleHisGlyThrAlaAspAepAenValHisPhe 707
QY 1217 GCACATACCAAGTATTTACTGAGTGTGTAGTGTGGGCTGGAAGCCATATATTTACAG 1276
D 708 GlnAenSerAlaGlnIleAlaLysAlaLeuValAsnAlaGlnValAspPheGlnAlaMet 727
QY 1277 ATCTATCTCAGGAGACACACAGCATAAGAGTTCTCTGAATCGGGA-----GAACAT 1327
```

us-10-825-632-6.n2p.rapbn

Mon May 8 13:41:56 2006

```
Db      728 TApTyrSerAspGlnAsnHisGlyLeu-----SerGlyLeuSerThrAsnHis 743
Qy      1328 TATGAACCTGCATCTTTTGGCACTACCTTCAAGAA 1360
Db      744 LeuTyrThrHisMetThrHisPheLeuLysGln 754
```

Search completed: May 2, 2006, 03:10:21
Job time : 54.8941 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 4, 2006, 02:53:29 ; Search time 5398 Seconds
(without alignments)
11404.501 Million cell updates/sec

Title: us-10-825-632-8
Perfect score: 1083
Sequence: 1 ggaagaagatgccagatcag.....tatgatacaggatacacgga 1083

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 2842172563 residues
Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba:*
2: gb_in:*
3: gb_env:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pr:*
9: gb_to:*
10: gb_sts:*
11: gb_sy:*
12: gb_un:*
13: gb_vi:*
14: gb_htg:*
15: gb_pl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1083	100.0	1083	6 AR651455	AR651455 Sequence
2	1083	100.0	1083	8 AF221637	AF221637 Homo sapi
3	1055.8	97.5	4309	6 AR631284	AR631284 Sequence
4	1055.8	97.5	4309	6 AX608737	AX608737 Sequence
5	911	84.1	2098	6 CQ721272	CQ721272 Sequence
6	791.2	73.1	2649	6 AX354795	AX354795 Sequence
7	791.2	73.1	2797	6 AR448400	AR448400 Sequence
8	791.2	73.1	3030	8 AY354202	AY354202 Homo sapi
9	791.2	73.1	3106	6 AX342633	AX342633 Sequence
10	791.2	73.1	3120	6 AR651452	AR651452 Sequence
11	791.2	73.1	3127	8 AF221634	AF221634 Homo sapi
12	791.2	73.1	3143	6 AX354793	AX354793 Sequence
13	791.2	73.1	4535	8 BC040203	BC040203 Homo sapi
14	789.6	72.9	2656	8 AY172659	AY172659 Homo sapi
15	789.6	72.9	2671	6 AR631278	AR631278 Sequence
16	789.6	72.9	2671	6 AX608725	AX608725 Sequence
17	779.2	71.9	4829	6 AR631283	AR631283 Sequence
18	779.2	71.9	4829	6 AX608735	AX608735 Sequence

19	669.6	61.8	4799	9 BC043124	BC043124 Mus muscu
20	669.6	61.8	4799	9 BC059222	BC059222 Mus muscu
21	606.6	56.0	753	6 AX524935	AX524935 Sequence
22	581.8	53.7	2510	6 AX338497	AX338497 Sequence
23	581.2	53.7	3125	8 BC030688	BC030688 Homo sapi
24	579.6	53.5	2161	6 BD157001	BD157001 Primer fo
25	579.6	53.5	2161	6 AX878058	AX878058 Sequence
26	579.6	53.5	2161	8 AK027826	AK027826 Homo sapi
27	549.2	50.7	1449	5 CR353491	CR353491 Gallus ga
28	497.2	45.9	4685	6 AR631288	AR631288 Sequence
29	497.2	45.9	4685	6 AX608745	AX608745 Sequence
30	479.2	44.2	4676	6 AR631287	AR631287 Sequence
31	479.2	44.2	4676	6 AX608743	AX608743 Sequence
32	459	42.4	1197	8 AR651453	AR651453 Sequence
33	459	42.4	1197	8 AF221635	AF221635 Homo sapi
34	434.2	40.1	1669	6 AR651454	AR651454 Sequence
35	434.2	40.1	1669	8 AF221636	AF221636 Homo sapi
36	434.2	40.1	2688	6 AX405771	AX405771 Sequence
37	434.2	40.1	2778	8 AK000290	AK000290 Homo sapi
38	434.2	40.1	2842	6 AX405770	AX405770 Sequence
39	428.2	39.5	4523	6 AR631281	AR631281 Sequence
40	428.2	39.5	4523	6 AX608731	AX608731 Sequence
41	385	35.5	183341	14 AC161443	AC161443 Mus muscu
42	347.4	32.1	2676	6 AX524928	AX524928 Sequence
43	344.2	31.8	2546	8 BC000970	BC000970 Homo sapi
44	344.2	31.8	2592	8 AY374518	AY374518 Homo sapi
45	344.2	31.8	2602	8 AY172660	AY172660 Homo sapi

ALIGNMENTS

RESULT 1
AR651455 AR651455 1083 bp DNA linear PAT 20-APR-2005
LOCUS Sequence 8 from patent US 6881564.
DEFINITION AR651455
ACCESSION AR651455
VERSION AR651455.1 GI:62795941
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1083)
AUTHORS Abbott, C.A. and Gorrell, M.D.
TITLE Dipeptidyl peptidases
JOURNAL Patent: US 6881564-A 8 19-APR-2005;
The University of Sydney; Sydney;
AUX;
FEATURES
source Location/Qualifiers
1..1083 /organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match	100.0%;	Score 1083;	DB 6;	Length 1083;
Best Local Similarity	100.0%;	Pred. No. 2.4e-256;		
Matches 1083;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	CGAAGAAGATCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAGAGAATTGTGATAG	60	
Db	1	CGAAGAAGATCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAGAGAATTGTGATAG	60	
QY	61	ATATTCTGGCTATTGGTGGTGTCCAAAGCTGAAACACTCCCGAGTGGTAAATTTCT	120	
Db	61	ATATTCTGGCTATTGGTGGTGTCCAAAGCTGAAACACTCCCGAGTGGTAAATTTCT	120	
QY	121	TAGAATTCCTATGAAGAAATGATGAATCTGAGGTGGAAATTTATTCATGTTACATCCCC	180	
Db	121	TAGAATTCCTATGAAGAAATGATGAATCTGAGGTGGAAATTTATTCATGTTACATCCCC	180	
QY	181	TATGTTGGAAACAAGGAGGCGAGATTCCGTTATCTCTTAAACAGGTACGCAATCC	240	
Db	181	TATGTTGGAAACAAGGAGGCGAGATTCCGTTATCTCTTAAACAGGTACGCAATCC	240	

QY	241	TAAAGTCACTTTTAAAGATGTCAGAAATAATGATTGATGCTGAAGGAAGATCATAGTTGA	300
Db	241	TAAAGTCACTTTTAAAGATGTCAGAAATAATGATTGATGCTGAAGGAAGATCATAGTTGA	300
QY	301	TGAAGTCAGAAGCGTGGTATATTTTGAAGCACCRAAGACTCCCTTTAGAGCATCACTT	360
Db	301	TGAAGTCAGAAGCGTGGTATATTTTGAAGCACCRAAGACTCCCTTTAGAGCATCACTT	360
QY	361	GTAGCTAGTCACTTACCTTAATCTGGAGAGGTGACAAGCTGACTGACCGTGGTACTC	420
Db	361	GTAGCTAGTCACTTACCTTAATCTGGAGAGGTGACAAGCTGACTGACCGTGGTACTC	420
QY	421	ACATTCCTGCTGATCAGTCAGTCAGCTGCTCTTTTAAAGTAAGTATAGTAAACAGAA	480
Db	421	ACATTCCTGCTGATCAGTCAGTCAGCTGCTCTTTTAAAGTAAGTATAGTAAACAGAA	480
QY	481	GAATCCACATGTGTGCTCTTTTACAGCTATCAAGTCTGGAAGATGACCCAACTTCAAA	540
Db	481	GAATCCACATGTGTGCTCTTTTACAGCTATCAAGTCTGGAAGATGACCCAACTTCAAA	540
QY	541	AACAAAGAAATTTTGGGCCACCAATTTTGGATTACAGAGTCTCTTCTGCTGACTATCTCC	600
Db	541	AACAAAGAAATTTTGGGCCACCAATTTTGGATTACAGAGTCTCTTCTGCTGACTATCTCC	600
QY	601	TCCAGAAATTTTCTCTTTTGAAGTACTACTGGAATTTACATGTATGGATGCTCTACAA	660
Db	601	TCCAGAAATTTTCTCTTTTGAAGTACTACTGGAATTTACATGTATGGATGCTCTACAA	660
QY	661	GCCTCATGATCATCAGCTTGGAAAGAAATATCTACTGTGCTGTTTCATATATGTTGGTCC	720
Db	661	GCCTCATGATCATCAGCTTGGAAAGAAATATCTACTGTGCTGTTTCATATATGTTGGTCC	720
QY	721	TCAGGTGCAGTGTGTGAATAATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCTT	780
Db	721	TCAGGTGCAGTGTGTGAATAATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCTT	780
QY	781	AGCCTCTCTAGTGTATGTGTTGCTAGTATGACACACAGGGGATCCTGTACCGAGGCT	840
Db	781	AGCCTCTCTAGTGTATGTGTTGCTAGTATGACACACAGGGGATCCTGTACCGAGGCT	840
QY	841	TAAATTTGAAGGCGCTTTAAATATATAAATGGTCAATAGAAATACAGATCAGGTGGA	900
Db	841	TAAATTTGAAGGCGCTTTAAATATATAAATGGTCAATAGAAATACAGATCAGGTGGA	900
QY	901	AGGACTCCATATCTAGCTTCTCGATGATGATTTCAATTCATAGATCGTGGGCATCCA	960
Db	901	AGGACTCCATATCTAGCTTCTCGATGATGATTTCAATTCATAGATCGTGGGCATCCA	960
QY	961	CGGCTGGTCTATGAGGATACCTCTCCCTGATGGCATTATGAGAGGTCAGATATCTT	1020
Db	961	CGGCTGGTCTATGAGGATACCTCTCCCTGATGGCATTATGAGAGGTCAGATATCTT	1020
QY	1021	CAGGTTGCTATTTGCTGGGGCCCCAGTCACTCTGFGAATCTTCTATGATACAGGATACAC	1080
Db	1021	CAGGTTGCTATTTGCTGGGGCCCCAGTCACTCTGFGAATCTTCTATGATACAGGATACAC	1080
QY	1081	GGA 1083	
Db	1081	GGA 1083	

RESULT 2
AF221637
LOCUS
DEFINITION Homo sapiens dipeptidyl peptidase 8 (DPP8) mRNA, partial cds,
alternatively spliced.
ACCESSION AF221637
VERSION AF221637.1 GI:11095193
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE AUTHORS	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominae; Homo. 1 (bases 1 to 1083) Abbott,C.A., Yu,D.M., Woollatt,E., Sutherland,G.R., McCaughan,G.W. and Gorrell,W.D. Cloning, expression and chromosomal localization of a novel human dipeptidyl peptidase (DPP) IV homolog, DPP8 Eur. J. Biochem. 267 (20), 6140-6150 (2000)
TITLE	
JOURNAL PUBMED	11012666
REFERENCE AUTHORS	2 (bases 1 to 1083) Abbott,C.A., Yu,D., McCaughan,G.W. and Gorrell,M.D. Direct Submission Submitted (06-JAN-2000) A.W. Morrow Gastroenterology and Liver Centre, Centenary Institute of Cell Biology and Cancer Medicine, Locked Bag No.6, Newtown, Sydney, NSW 2042, Australia
JOURNAL TITLE	Location/Qualifiers
FEATURES	1..1083 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /chromosome="15" /map="15q22" /clone="stPBMCd3-3-1" /cell_type="PHA stimulated PBMC" <1..>1083 /genes="DPP8" <1..>1083 /genes="DPP8" /notes="serine protease, member of prolyl oligopeptidase S9b family belonging to enzyme clan SC; additional splice variants of this mRNA are found in GenBank Accession Numbers AF221634, AF221635 and AF221636" /codon_start=2 /product="dipeptidyl peptidase 8" /protein_id="AA029769.1" /db_xref="GI:11095194" /translation="BEDARSAGVATFVLQEPDRYGVWCPKAEPTSPGKILRILY EENDESEIIHVTSPMLETRRADSFYKPTGTANPKVTFKMSIMIDAERIIIVDEV KRIVYEGKDSFLEHLXVSVNPGVETRLTDRGYSHSCCIQHQCDFFIFSKYSNQK RPSVLSYKLSFSEDDPTCKTFEWAATLDSAGLPDYPPEIIFSEFTTFTLYGML YKPHDLPQPKYPTVLFIYGGPQVLVNNRFGVKYFRLNLASLGVVVVVDNRGSC HRLKEPEGAPKYMKGQIEIDDQVQLASRYDFIDLRVGIHGSYGGYLSLMLM QRSIDIFRVAIAGPVTLMIFYDTGYT"
gene	
CDS	
ORIGIN	
Query Match	100.0%; Score 1083; DB 8; Length 1083;
Best Local Similarity	100.0%; Pred. No. 2.4e-256;
Matches 1083; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	1 GGAAGAAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAAGAAGAATTGTAG 60
Db	1 GGAAGAAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAAGAAGAATTGTAG 60
QY	61 ATATTCTGGCTATTGGTGGTGTCCAAAAGCTGAAAACAACTCCAGTGGTGGTAAAAATCT 120
Db	61 ATATTCTGGCTATTGGTGGTGTCCAAAAGCTGAAAACAACTCCAGTGGTGGTAAAAATCT 120
QY	121 TAGAATTTCTATATGAAGAAAATGATGAATCTGAGGTGGAATTTATTCATGTTACATCCC 180
Db	121 TAGAATTTCTATATGAAGAAAATGATGAATCTGAGGTGGAATTTATTCATGTTACATCCC 180
QY	181 TATGTTTGGAAAACAGGAGGCGAGATTTCCTGGTTATCTTAAACAGGTACAGCAATCC 240
Db	181 TATGTTTGGAAAACAGGAGGCGAGATTTCCTGGTTATCTTAAACAGGTACAGCAATCC 240
QY	241 TAAAGTCACCTTTTAAAGATGTCAGAAATATGATGATGCTCAGAGGAGGATCATAGTTGA 300
Db	241 TAAAGTCACCTTTTAAAGATGTCAGAAATATGATGATGCTCAGAGGAGGATCATAGTTGA 300
QY	301 TGAAGTCAGAGGCTGGTATATTTTGAAGGACCAAGAGATCCCTTTAGAGCATCACTT 360
Db	301 TGAAGTCAGAGGCTGGTATATTTTGAAGGACCAAGAGATCCCTTTAGAGCATCACTT 360


```
QY 1019 TTCAGGTTGCTATTGCTGGGGCCCGCAGTCACCTCTGTGGATCTTCTATGATACAGGATAC 1078
Db 1961 TTCAGGTTGCTATTGCTGGGGCCCGCAGTCACCTCTGTGGATCTTCTATGATACAGGATAC 2020
QY 1079 ACGGA 1083
Db 2021 ACGGA 2025

RESULT 4
AX608737
LOCUS AX608737 4309 bp DNA linear PAT 17-FEB-2003
DEFINITION Sequence 14 from Patent WO0231134.
ACCESSION AX608737
VERSION AX608737.1 GI:28404304
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1
REFERENCE Qi, S., Akinsanya, K.O., Riviere, P.J. and Junien, J.L.
AUTHORS Novel serine protease genes related to dppiv
TITLE Patent: WO 0231134-A 14 18-APR-2002;
JOURNAL Ferring BV (NL)
FEATURES
source Location/Qualifiers
1 4309
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 97.5%; Score 1055.8; DB 6; Length 4309;
Best Local Similarity 99.5%; Pred. No. 1.1e-249;
Matches 1080; Conservative 0; Mismatches 2; Indels 3; Gaps 2;

QY 1 GGAAGAAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAAAGAAGATTGTATAG 60
Db 942 GGAAGAAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAAAGAAGATTGTATAG 1001
QY 61 ATATCTCGCTATTGGTGGTGTCCAAAGCTGAACCAACTCCCACTGGTGGTGAATTTCT 120
Db 1002 ATATCTCGCTATTGGTGGTGTCCAAAGCTGAACCAACTCCCACTGGTGGTGAATTTCT 1061
QY 121 TAGAATTTCTATATGAAGAAATGATGAATCTGAGGTGGAAATTTATTCATGTATACATCCC 180
Db 1062 TAGAATTTCTATATGAAGAAATGATGAATCTGAGGTGGAAATTTATTCATGTATACATCCC 1121
QY 181 TATGTTGGAACCAAGAGGCGCAGATTCATTCGTTATCTTAAACAGGTACAGCAAAATCC 240
Db 1122 TATGTTGGAACCAAGAGGCGCAGATTCATTCGTTATCTTAAACAGGTACAGCAAAATCC 1181
QY 241 TAAAGTCACCTTTTAAGATGTCAGAAATATGATTGATGCTGAAGGAAGGATCATAGTTGA 300
Db 1182 TAAAGTCACCTTTTAAGATGTCAGAAATATGATTGATGCTGAAGGAAGGATCATAGTTGA 1240
QY 301 TGAAGTCAGAGGCTGTATATTTTGAAGCCCAAGACTCCCTTTAGAGCATCACCT 360
Db 1241 TGAAGTCAGAGGCTGTATATTTTGAAGCCCAAGACTCCCTTTAGAGCATCACCT 1300
QY 361 GTAGTGTACGTTACCTTAATCTCGAGAGGTGACAGGCTGATCGTCCGCTGGTACTC 420
Db 1301 GTAGTGTACGTTACCTTAATCTCGAGAGGTGACAGGCTGATCGTCCGCTGGTACTC 1360
QY 421 ACATCTTGCTGCATCAGTCAGCACTGTGACTCTTTTAAAGTATAGTATAGTACACGAA 480
Db 1361 ACATCTTGCTGCATCAGTCAGCACTGTGACTCTTTTAAAGTATAGTATAGTATAGTACACGAA 1420
QY 481 GAATCCCACTGTGTGCTCTTTTACAGGTATCAAGTCTCGAAGATGACCCAACTTGCAG 540
Db 1421 GAATCCCACTGTGTGCTCTTTTACAGGTATCAAGTCTCGAAGATGACCCAACTTGCAG 1480
```

```
QY 541 AACAAAGGAATTTTGGGCCACCAATTTTGGATTCCAGCAGGTCTCTTCTGACTATACTCC 600
Db 1481 AACAAAGGAATTTTGGGCCACCAATTTTGGATTCCAGCAGGTCTCTTCTGACTATACTCC 1540
QY 601 TCCAGAAATTTTCTCTTTTGAAGTACTACTGATTTACATTTGATGGATGCTCTACAA 660
Db 1541 TCCAGAAATTTTCTCTTTTGAAGTACTACTGATTTACATTTGATGGATGCTCTACAA 1600
QY 661 GCCTCATGATCTACAGCCTGGAAGAAATATCTACTGCTGCTGTTTCATATATGGTGG--T 718
Db 1601 GCCTCATGATCTACAGCCTGGAAGAAATATCTACTGCTGCTGTTTCATATATGGTGGTCT 1660
QY 719 COTCAGGTGCAAGTTGGTGAATATCGGTTTAAAGAGGAGTCAAGTATTTCCGCTTGAATACC 778
Db 1661 COTCAGGTGCAAGTTGGTGAATATCGGTTTAAAGAGGAGTCAAGTATTTCCGCTTGAATACC 1720
QY 779 CTAGCCTCTCTAGGTTATGTGGTTGTAGTGATAGACAAACAGGGGATCTCTGCCAGGG 838
Db 1721 CTAGCCTCTCTAGGTTATGTGGTTGTAGTGATAGACAAACAGGGGATCTCTGCCAGGG 1780
QY 839 CTTAAATTTGAAGGCGCTTTTAAATATAAATGGGTCAAATAGAAATTCACGATCAGGTG 898
Db 1781 CTTAAATTTGAAGGCGCTTTTAAATATAAATGGGTCAAATAGAAATTCACGATCAGGTG 1840
QY 899 GAAGGACTCCCAATATCTAGCTTCTCGATATGATTTTCATTTAGATCGTGTGGCATC 958
Db 1841 GAAGGACTCCCAATATCTAGCTTCTCGATATGATTTTCATTTAGATCGTGTGGCATC 1900
QY 959 CACGCTGGTCTATGAGGATACCTCTCCCTGATGGCATTAATGCAGAGTTCAGATATC 1018
Db 1901 CACGCTGGTCTATGAGGATACCTCTCCCTGATGGCATTAATGCAGAGTTCAGATATC 1960
QY 1019 TTCAGGTTGCTATTGCTGGGGCCCGCAGTCACCTCTGTGGATCTTCTATGATACAGGATAC 1078
Db 1961 TTCAGGTTGCTATTGCTGGGGCCCGCAGTCACCTCTGTGGATCTTCTATGATACAGGATAC 2020
QY 1079 ACGGA 1083
Db 2021 ACGGA 2025

RESULT 5
CO721272 2098 bp DNA linear PAT 03-FEB-2004
LOCUS Sequence 7206 from Patent WO02068579.
DEFINITION CO721272
ACCESSION CO721272
VERSION CO721272.1 GI:42282129
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1
REFERENCE Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
AUTHORS Kites, such as nucleic acid arrays, comprising a majority of
TITLE humanexons or transcripts, for detecting expression and other uses
thereof
JOURNAL Patent: WO 02068579-A 7206 06-SEP-2002;
PE Corporation (NY) (US)
FEATURES
source Location/Qualifiers
1..2098
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 84.1%; Score 911; DB 6; Length 2098;
Best Local Similarity 88.9%; Pred. No. 6.2e-214;
Matches 1048; Conservative 0; Mismatches 35; Indels 96; Gaps 2;

QY 1 GGAAGAAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAAAGAAGAAATTTGATAG 60
```

Db	835	GGAGAGATCCAGATCAGCTGGAGTCGCTACTCTTTGTTCTCCAGAGAAATTTGATAG	894
Qy	61	ATAATTCGGCTATTGGTGGTGTCCAAAGCTGAAACAACTCCAGTGGTGTAATAATCT	120
Db	895	ATAATTCGGCTATTGGTGGTGTCCAAAGCTGAAACAACTCCAGTGGTGTAATAATCT	954
Qy	121	TAGAAATTCATATGAAGAATGATGAATCTGAGGTGGAAATTAATTCATGTACATCCC	180
Db	955	TAGAAATTCATATGAAGAATGATGAATCTGAGGTGGAAATTAATTCATGTACATCCC	1014
Qy	181	TATGTTGGAACAAGGAGGCGAGATTCATTCCTGTTATCTCTAAACAGGT-----	229
Db	1015	TATGTTGGAACAAGGAGGCGAGATTCATTCCTGTTATCTCTAAACAGATCGCTCCAGAC	1074
Qy	230	-----	229
Db	1075	TCGCCACAGATAGTGTGTGATCTCACCTGAAATTAATTTATCCCGAGTAGAAGATGTTAT	1134
Qy	230	-----ACAGCAATCTAAAGTCATCTTTAAGATGTCAGAAATAAATGATGATGC	279
Db	1135	GGAAAGGACAGAGATCATTTGAGTCAGTCGCTGATCTGTGACGCCACTAAATTAATCTATGA	1194
Qy	280	TGAAGGAGGATCAT-----AGTTGATGAAGTCAGAGCGCTGGTATATT	324
Db	1195	AGAAACAACAGACATCTGGATAAATATCCAAAGTTGATGAAGTCAGAGCGCTGGTATATT	1254
Qy	325	TGAAGGACCAAGACCTCCCTTTAGAGCATCACCTGTACGTAGTCAGTTACGTAAATCC	384
Db	1255	TGAAGGACCAAGACCTCCCTTTAGAGCATCACCTGTACGTAGTCAGTTACGTAAATCC	1314
Qy	385	TGGAGAGGTGACAAAGGCTGACTGACCGTGCTACTCACATTTCTTGCTGCATCAGTCAGCA	444
Db	1315	TGGAGAGGTGACAAAGGCTGACTGACCGTGCTACTCACATTTCTTGCTGCATCAGTCAGCA	1374
Qy	445	CTGTGACTCTTTTAAAGTAAATAGTAAACCAAGAAATCCACACTGTGTGTCCTCTTTA	504
Db	1375	CTGTGACTCTTTTAAAGTAAATAGTAAACCAAGAAATCCACACTGTGTGTCCTCTTTA	1434
Qy	505	CAAGCTATCAAGTCTGAGAGTACCCTGCAAAAACAAAGGAATTTTGGGCCACCAT	564
Db	1435	CAAGCTATCAAGTCTGAGAGTACCCTGCAAAAACAAAGGAATTTTGGGCCACCAT	1494
Qy	565	TTTGGATTACAGAGTCTCTTCTGCTACTATCTCCCTCAGAAATTTCTCTTTTGAAG	624
Db	1495	TTTGGATTACAGAGTCTCTTCTGCTACTATCTCCCTCAGAAATTTCTCTTTTGAAG	1554
Qy	625	TACTACTGATTTTACATTTGATGGATGCTCTACAAGCTCATGATCTACAGCCTGGAAA	684
Db	1555	TACTACTGATTTTACATTTGATGGATGCTCTACAAGCTCATGATCTACAGCCTGGAAA	1614
Qy	685	GAAATATCCTACTGTGCTTCAATATGTTGGTGTCTCAGGTGCGAGTTGGTGAATAATCG	744
Db	1615	GAAATATCCTACTGTGCTTCAATATGTTGGTGTCTCAGGTGCGAGTTGGTGAATAATCG	1674
Qy	745	GTTTAAAGAGTCAAGTATTTCCGCTTGAATACCCCTAGCCTCTCTAGGTTATGTGGTTGT	804
Db	1675	GTTTAAAGAGTCAAGTATTTCCGCTTGAATACCCCTAGCCTCTCTAGGTTATGTGGTTGT	1734
Qy	805	AGTGATAGACAACAGGGATCCTGTGACCGAGGGCTTAAATTTGAAGGGCCCTTTAAATA	864
Db	1735	AGTGATAGACAACAGGGATCCTGTGACCGAGGGCTTAAATTTGAAGGGCCCTTTAAATA	1794
Qy	865	TAAATGGGTCAATAGAAATTCAGATCAGGTGGAGGAGCTCCATATCTAGCTTCG	924
Db	1795	TAAATGGGTCAATAGAAATTCAGATCAGGTGGAGGAGCTCCATATCTAGCTTCG	1854
Qy	925	ATATGATTTTCAATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT	984
Db	1855	ATATGATTTTCAATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT	1914
Qy	985	CTCCCTGATGTCATTAATGCAGAGGTGATATATCTTCAGGGTTGCTTATTCCTGGGGCCCC	1044
Db	1915	CTCCCTGATGTCATTAATGCAGAGGTGATATATCTTCAGGGTTGCTTATTCCTGGGGCCCC	1974
Qy	1045	AGTCACTCTGTGGATCTTCTATGATACAGATACACGGA	1083
Db	1975	AGTCACTCTGTGGATCTTCTATGATACAGATACACGGA	2013
RESULT 6			
AX354795			
LOCUS	AX354795	2649 bp	DNA linear PAT 06-FEB-2002
DEFINITION	Sequence 3 from Patent WO0179473.		
ACCESSION	AX354795		
VERSION	AX354795.1	GI:18619528	
KEYWORDS	Homo sapiens (human)		
SOURCE	Homo sapiens		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1		
AUTHORS	Meyers, R.A. and Williamson, M.		
TITLE	21953, a human prolyl oligopeptidase family member and uses thereof		
JOURNAL	Patent: WO 0179473-A 3 25-OCT-2001;		
	Millennium Pharmaceuticals, Inc. (US)		
FEATURES	Location/Qualifiers		
source	1..2649		
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
ORIGIN			
Query Match	73.1%; Score 791.2; DB 6; Length 2649;		
Best Local Similarity	99.0%; Pred. No. 2.2e-184;		
Matches	796; Conservative 0; Mismatches 8; Indels 0; Gaps 0;		
Qy	280	TGAAGGAAGGATCATAGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGGCACCAAGA	339
Db	1527	TGGATCTAATATCCAAAGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGGCACCAAGA	1586
Qy	340	CTCCCTTTAGAGCATCACCTGTACGTAGTCAGTTAGCTAAATCCTGGAGAGGTGACAAG	399
Db	1587	CTCCCTTTAGAGCATCACCTGTAGTCAGTTAGCTAAATCCTGGAGAGGTGACAAG	1646
Qy	400	GCTGACTCAGCGTGCCTACTCACATTTCTGTCATCAGTCAGCACCTGTGACTTCTTTAT	459
Db	1647	GCTGACTCAGCGTGCCTACTCACATTTCTGTCATCAGTCAGCACCTGTGACTTCTTTAT	1706
Qy	460	AACTAAGTATAGTAAACCAAGAAATCCACACTGTGTGCTCCCTTTACAAGCTATCAAGTCC	519
Db	1707	AACTAAGTATAGTAAACCAAGAAATCCACACTGTGTGCTCCCTTTACAAGCTATCAAGTCC	1766
Qy	520	TGAAGATCACCAACTTGCAGAAACAAAGGAATTTTGGGCCACCATTTTGGATTACAGAG	579
Db	1767	TGAAGATCACCAACTTGCAGAAACAAAGGAATTTTGGGCCACCATTTTGGATTACAGAG	1826
Qy	580	TCCTCTTCTGACTATCTCTCAGAAATTTTCTCTTTTGAAGTACTACTGGATTAC	639
Db	1827	TCCTCTTCTGACTATCTCTCAGAAATTTTCTCTTTTGAAGTACTACTGGATTAC	1886
Qy	640	ATTGTTATGGATGCTCTTACAGCCTCATGATCTACAGCCTGGAAAGAAATATCTACTGT	699
Db	1887	ATTGTTATGGATGCTCTTACAGCCTCATGATCTACAGCCTGGAAAGAAATATCTACTGT	1946
Qy	700	GCTGTTTATATATATGTTGCTCTCAGGTGCTAGTTGGTGAATAATCGGTTTAAAGGAGTCAA	759
Db	1947	GCTGTTTATATATATGTTGCTCTCAGGTGCTAGTTGGTGAATAATCGGTTTAAAGGAGTCAA	2006
Qy	760	GTATTTCCGCTTGAATACCCCTAGCCTCTCTAGGTTATGTGTTGTAGTGATGACACACAG	819
Db	2007	GTATTTCCGCTTGAATACCCCTAGCCTCTCTAGGTTATGTGTTGTAGTGATGACACACAG	2066
Qy	820	GGGATCCTGTCCAGAGGGCTTAAATTTGAAGGCGCTTTTAAATATATAAATGGGTCAAT	879

Db 2067 GGGATCCTGTACCAGGGCTTAAATTTGAAGGGCCCTTTAAATATATAAAATGGGTCAAT 2126
QY 880 AGAATTCACGATCAGGTGGAGGACTCCGAATATCTAGCTTCTCGATATGATTTCAATGA 939
Db 2127 AGAATTCACGATCAGGTGGAGGACTCCGAATATCTAGCTTCTCGATATGATTTCAATGA 2186
QY 940 CTTAGATCGTGTGGGATCCACGGCTGGTCTCTATGAGGATACCTCTCCCTGATGGCATTT 999
Db 2187 CTTAGATCGTGTGGGATCCACGGCTGGTCTCTATGAGGATACCTCTCCCTGATGGCATTT 2246
QY 1000 AATGAGAGGTTCAGATATCTTCAGGGTTGCTATTGCTGGGGCCCGAGTCACCTCTGTGGAT 1059
Db 2247 AATGAGAGGTTCAGATATCTTCAGGGTTGCTATTGCTGGGGCCCGAGTCACCTCTGTGGAT 2306
QY 1060 CTTCTATGATACAGGATACACGGA 1083
Db 2307 CTTCTATGATACAGGATACACGGA 2330

RESULT 7
AR448400
LOCUS AR448400 2797 bp DNA linear PAT 20-FEB-2004
DEFINITION Sequence 1103 from patent US 6673549.
ACCESSION AR448400
VERSION AR448400.1 GI:42676724
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.

REFERENCE
1 (bases 1 to 2797)
Furness, L.M. and Buchbinder, J.L.
Genes expressed in C3A liver cell cultures treated with steroids
Patent: US 6673549-A, 1103 06-JAN-2004;
Incyte Corporation; Palo Alto, CA
LOCATION/Qualifiers
1..2797
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 73.1%; Score 791.2; DB 6; Length 2797;
Best Local Similarity 99.0%; Pred. No. 2.2e-184;
Matches 796; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 280 TGAAGGAAGGATCATAGTTGATGAAGTCAAGGGCTGGTATATTTTGAAGGCCACCAAGA 339
Db 1411 TGGATCTAATATCCAAGTTGATGAAGTCAAGGGCTGGTATATTTTGAAGGCCACCAAGA 1470
QY 340 CTCCTCTTTAGAGATCACCTGTACGTAGTACGTAAATCCTGGAGAGGTGACAAG 399
Db 1471 CTCCTCTTTAGAGATCACCTGTACGTAGTACGTAAATCCTGGAGAGGTGACAAG 1530
QY 400 GTCGACTGACCGTGGCTACTACATTTCTTGCTGCATCAGTCAGCACTGTGACTTCTTTAT 459
Db 1531 GTCGACTGACCGTGGCTACTACATTTCTTGCTGCATCAGTCAGCACTGTGACTTCTTTAT 1590
QY 460 AAGTAAGTATAGTAACCAAGAAATCCACACTGTGTGCTCCCTTACAGCTATCAAGTCC 519
Db 1591 AAGTAAGTATAGTAACCAAGAAATCCACACTGTGTGCTCCCTTACAGCTATCAAGTCC 1650
QY 520 TGAAGTACCGCCACTGCGAACAAGAAATTTGGGCCACCATTTGGATTCAGCAGG 579
Db 1651 TGAAGTACCGCCACTGCGAACAAGAAATTTGGGCCACCATTTGGATTCAGCAGG 1710
QY 580 TCCTCTCTCTGACTATCTCTCCAGAAATTTCTCTTTTGAAGTACTACTGGATTAC 639
Db 1711 TCCTCTCTCTGACTATCTCTCCAGAAATTTCTCTTTTGAAGTACTACTGGATTAC 1770
QY 640 ATTGTATGGATGCTCTACAAGGCTCATGATCTACAGGCTGGAAGAATAATCTACTGT 699
Db 1771 ATTGTATGGATGCTCTACAAGGCTCATGATCTACAGGCTGGAAGAATAATCTACTGT 1830
QY 700 GCTGTTTCATATATGGTGGTCCCTCAGGTGAGTTGGTGAATATATCGGTTTAAAGGAGTCAA 759

Db 1831 GCTGTTTCATATATGGTGGTCCCTCAGGTGCAGTTTGGTGAATATCGGTTTAAAGGAGTCAA 1890
QY 760 GTATTTCCGCTTGAATACCTTAGCCTCTTAGGTTATGCTGTTGTTAGTGATGATGACAACAG 819
Db 1891 GTATTTCCGCTTGAATACCTTAGCCTCTTAGGTTATGCTGTTGTTAGTGATGATGACAACAG 1950
QY 820 GGGATCCTGTACCAGAGGGCTTAAATTTGAAGGGCCCTTTAAATATATAAAATGGGTCAAT 879
Db 1951 GGGATCCTGTACCAGAGGGCTTAAATTTGAAGGGCCCTTTAAATATATAAAATGGGTCAAT 2010
QY 880 AGAATTCACGATCAGGTGGAGGACTCCGAATATCTAGCTTCTCGATATGATTTCAATGA 939
Db 2011 AGAATTCACGATCAGGTGGAGGACTCCGAATATCTAGCTTCTCGATATGATTTCAATGA 2070
QY 940 CTTAGATCGTGTGGGATCCACGGCTGGTCTCTATGAGGATACCTCTCCCTGATGGCATTT 999
Db 2071 CTTAGATCGTGTGGGATCCACGGCTGGTCTCTATGAGGATACCTCTCCCTGATGGCATTT 2130
QY 1000 AATGAGAGGTTCAGATATCTTCAGGGTTGCTATTGCTGGGGCCCGAGTCACCTCTGTGGAT 1059
Db 2131 AATGAGAGGTTCAGATATCTTCAGGGTTGCTATTGCTGGGGCCCGAGTCACCTCTGTGGAT 2190
QY 1060 CTTCTATGATACAGGATACACGGA 1083
Db 2191 CTTCTATGATACAGGATACACGGA 2214

RESULT 8
AY354202
LOCUS AY354202 3030 bp mRNA linear PRI 02-SEP-2003
DEFINITION Homo sapiens dipeptidyl peptidase 8 isoform 3 (DPP8) mRNA, complete cds, alternatively spliced.
ACCESSION AY354202
VERSION AY354202.1 GI:34329351
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 3030)
Sha, J.H., Zhou, Z.M. and Li, J.M.
Direct Submission
Submitted (29-JUL-2003) Key Lab of Reproductive Medicine, Nanjing Medical University, 140 Han Zhong Road, Nanjing, Jiangsu 210029, China
LOCATION/Qualifiers
1..3030
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="15"
/tissue_type="testis"
1..3030
/gene="DPP8"
181..2857
/gene="DPP8"
/note="alternatively spliced"
/codon_start=1
/product="dipeptidyl peptidase 8 isoform 3"
/protein_id="AA063887.1"
/db_xref="GI:34329352"
/translation="MKRSEQMKIKSGKCNMAAEMETQLGVIEFETADCBENIESQD
RKLPEPVRYRSQKLKLLADTRKYHYMMAKPHDFMFKRNDPDGPHSDRIYLL
AMGSENRENTLFYSEIPKTNINRAALMLSWKPLDLFOATLDYGMYSREBLRKR
IGTVDASYDHYHQSGLTFQAGSGIYHVKGDPGQPTQQPLRNLVETSCFNIRMDP
KLCPADPDWTFATHSNDIWSNIVTREERLLTYVHNEANMEEDARSAGVATVLQEE
FDYSGVWCFKAEITPSSGKILRLIYEENDESEVEIHHVTSPLMELARADSPRYKET
GTANPKVTFKMSIEMIDAEGRIIDVDEKELIQFELFEGVEYIARAGWTEGKYAWS
ILDRSQTRLQIVLSIPELFIVEDDWMERQRLIESVPDSVTPPLIYIETTDWINIH
DIHFVPSQHEEIEIFAFSECKTFRHLKYITSLIKESKYKRSNGGLPSPDFKCP
KEEITATSGEWEVLGRHSGNIQVDEVRLVYFEGTKDSPLEHLLVYVSVNPGVEVRL

ORIGIN									
IFASECKTGFRHLXYKITSILKESKYRSGGLPAPSPDFKCPKEIEIAITSGWEVLGR									
HGSNIQDVSRVLVVEGKDSPLHLYVYVYVNPCEVTRLDRGVSHSCCISOHCD									
PFIISKVSNQKNPHCVSLYKLSPEDDPCTKKEFWAILDSAGPLDYPPEIFSPES									
TTGFTLYGLMYKPHDLQPKKYPTVFLFYGGPQVQLVNNRFGVKYFRNLTLASLGYV									
VVVIDNRGSGHRLKPEGAFFKMGQIBIDDOVEGLQYLASRYDFDLDRVGHWSY									
GGVLSLMALMORSIDIFRVAIAGAPVTLMIFYDTGYTRYMGHPDQNEQGYLGSVAMQ									
AEKFPSEPNRLLLHGLFDENVHFAHTSILLSFLVRAGKPYDLQIYQPERHSIRVPES									
GEHYELHLLHYLQENLGSRIAAKVI"									
ORIGIN									
Query Match 73.1%; Score 791.2; DB 8; Length 3127;									
Best Local Similarity 99.0%; Pred. No. 2.2e-184;									
Matches 796; Conservative 0; Mismatches 8; Indels 0; Gaps 0;									
QY	280	TGAAGGAGGATCATAGTTGATGAAGTCAGAAAGCGTGCTATATTTTGAAGGCCACCAAGA	339						
DB	1740	TGGATCTAATATCCAAGTTGATGAAGTCAGAAAGCGTGCTATATTTTGAAGGCCACCAAGA	1799						
QY	340	CTCCCTTTAGAGCATCACTGTACGTAGTCAGTTACGTAATCTCGAGAGGTGACAAG	399						
DB	1800	CTCCCTTTAGAGCATCACTGTACGTAGTCAGTTACGTAATCTCGAGAGGTGACAAG	1859						
QY	400	GCTGACTGACCGTGGCTACTACATCTTGCTGTCATCAGTCAGCAGCTGTGACTTCTTTAT	459						
DB	1860	GCTGACTGACCGTGGCTACTACATCTTGCTGTCATCAGTCAGCAGCTGTGACTTCTTTAT	1919						
QY	460	AAGTAAGTATAGTAACCAAGAAATCCACACTGTGTGCCCTTTTACAAGCTATCAAGTCC	519						
DB	1920	AAGTAAGTATAGTAACCAAGAAATCCACACTGTGTGCCCTTTTACAAGCTATCAAGTCC	1979						
QY	520	TGAAGTACCCCAACTTGCAAAACAAAGGAAATTTTGGGCCACCACTTTTGGATTGACAGG	579						
DB	1980	TGAAGTACCCCAACTTGCAAAACAAAGGAAATTTTGGGCCACCACTTTTGGATTGACAGG	2039						
QY	580	TCCTCTTCTGACTATATCTCTCAGAAATTTTCTCTTTGAAAGTACTACTGATTTAC	639						
DB	2040	TCCTCTTCTGACTATATCTCTCAGAAATTTTCTCTTTGAAAGTACTACTGATTTAC	2099						
QY	640	ATTGTATGGGATGCTTCAAGCCTCATGATCTACAGCTGGAAGAAATATCCTACTGT	699						
DB	2100	ATTGTATGGGATGCTTCAAGCCTCATGATCTACAGCTGGAAGAAATATCCTACTGT	2159						
QY	700	GCTGTTTCATATATGTTGCTCCTCAGTGCAGTTGGTGAATAATCGGTTTAAAGAGTCAA	759						
DB	2160	GCTGTTTCATATATGTTGCTCCTCAGTGCAGTTGGTGAATAATCGGTTTAAAGAGTCAA	2219						
QY	760	GTATTTCCGCTTGAATACCCCTAGCCTCTCTAGGTTATGTTGTTAGTATAGACACAG	819						
DB	2220	GTATTTCCGCTTGAATACCCCTAGCCTCTCTAGGTTATGTTGTTAGTATAGACACAG	2279						
QY	820	GGGATCTCTGACCCGAGGGCTTAAATTTGAAGGCGCCTTTAAATATAAAATGGGTCAAAT	879						
DB	2280	GGGATCTCTGACCCGAGGGCTTAAATTTGAAGGCGCCTTTAAATATAAAATGGGTCAAAT	2339						
QY	880	AGAAATGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTCAATGA	939						
DB	2340	AGAAATGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTCAATGA	2399						
QY	940	CTTAGATCGTGTGGGCATCCACGGCTGGTCTTATGGAGGATACCTCTCCCTGATGGCATT	999						
DB	2400	CTTAGATCGTGTGGGCATCCACGGCTGGTCTTATGGAGGATACCTCTCCCTGATGGCATT	2459						
QY	1000	AATGCAGAGGTTCAGATATCTTACAGGGTTGCTATTTGCTGGGGCCCCAGTCACTCTGTGGAT	1059						
DB	2460	AATGCAGAGGTTCAGATATCTTACAGGGTTGCTATTTGCTGGGGCCCCAGTCACTCTGTGGAT	2519						
QY	1060	CTTCTATGATACAGGATACACGGA 1083							
DB	2520	CTTCTATGATACAGGATACACGGA 2543							
RESULT 12									
AX354793									

LOCUS	AX354793	3143 bp	DNA	linear	PAT 06-FEB-2002
DEFINITION	Sequence 1 from Patent WO0179473.				
ACCESSION	AX354793				
VERSION	AX354793.1	GI:18619526			
KEYWORDS	Homo sapiens (human)				
SOURCE	Homo sapiens				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1				
AUTHORS	Meyers, R.A. and Williamson, M.				
TITLE	21953, a human prolyl oligopeptidase family member and uses thereof				
JOURNAL	Patent: WO 0179473-A 1 25-OCT-2001; Millennium Pharmaceuticals, Inc. (US)				
FEATURES	Location/Qualifiers				
source	1..3143				
	/organism="Homo sapiens"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:9606"				
CDS	229..2877				
	/note="unnamed protein product"				
	/codon_start=1				
	/protein_id="CAD22990.1"				
	/db_xref="GI:18619527"				
	/translation="MAAAWETBQLGVEIFETADCEENIESQDRPKLEPFYVERYSWQ LKLLADTRKHYWMAKPHDFMFKRNDPDGSHSDRIYYLAMSNGRENTLFYSEI PKTINRAAVLMSWKLPLDLFQATLDYGYMREBELLRERKRIQTVGLIASYDYHQSG TFIIFOAGSGIYHVKGQPGFTQOPLRNVLVETSCPNRMDPKLPADPWIAFHSHN DIWISNIVTREERRLLTYVHNELANNEEDARSAGVATFVLOEFDRYSGYMWCPKAEIT PSGKILRLIYBENDESEVEIHLVTSPLMETRRADSFRYPKGTANPKVTFKMESEIMI DABGRIIDVDELQPFELFEGVEYIARAGWTPGKYAWSILLDRSQRLQLVLIS PELFIPEDDVMERQLESVDSVTPLLIYEETDIIWINTHIDFHVFPQSHHEEIEF IFASECKTGFRHLXYKITSILKESKVRSSGGLPAPSPDFKCPKEIEIAITSGBEWELGR HGSNIQDVSRVLVVEGKDSPLHLYVYVYVNPCEVTRLDRGVSHSCCISOHCD TTGFTLLYGLMYKPHDLQPKKYPTVFLFYGGPQVQLVNNRFGVKYFRNLTLASLGYV VVIDNRGSGHRLKPEGAFFKMGQIBIDDOVEGLQYLASRYDFDLDRVGHWSY GGYLSLMALMORSIDIFRVAIAGAPVTLMIFYDTGYTRYMGHPDQNEQGYLGSVAMQ AEKFPSEPNRLLLHGLFDENVHFAHTSILLSFLVRAGKPYDLQIYQPERHSIRVPES GEHYELHLLHYLQENLGSRIALKVI"				
ORIGIN					
Query Match	73.1%; Score 791.2; DB 6; Length 3143;				
Best Local Similarity	99.0%; Pred. No. 2.2e-184;				
Matches 796; Conservative	0; Mismatches 8; Indels 0; Gaps 0;				
Qy	280	TGAAGGAGGATCATAGTTGATGAAGTCAGAAAGCGTGCTATATTTTGAAGGCCACCAAGA	339		
Db	1755	TGGATCTAATATCAAGTTGATGAAGTCAGAAAGCGTGCTATATTTTGAAGGCCACCAAGA	1814		
Qy	340	CTCCCTTTAGAGCATCACCTGTACGTAGTCAGTTACGTAATCTCGAGAGGTGACAAG	399		
Db	1815	CTCCCTTTAGAGCATCACCTGTACGTAGTCAGTTACGTAATCTCGAGAGGTGACAAG	1874		
Qy	400	GCTGACTGACCGTGGCTACTCAATCTTTGCTGTCATCAGTCAGCAGCTGTGACTTCTTTAT	459		
Db	1875	GCTGACTGACCGTGGCTACTCAATCTTTGCTGTCATCAGTCAGCAGCTGTGACTTCTTTAT	1934		
Qy	460	AAGTAAGTATAGTAACCAAGAAATTTTGGGCCACCACTTTTGGATTGACAGTC	519		
Db	1935	AAGTAAGTATAGTAACCAAGAAATTTTGGGCCACCACTTTTGGATTGACAGTC	1994		
Qy	520	TGAAGTACCCCAACTTGCAAAACAAAGGAAATTTTGGGCCACCACTTTTGGATTGACAGCAGG	579		
Db	1995	TGAAGTACCCCAACTTGCAAAACAAAGGAAATTTTGGGCCACCACTTTTGGATTGACAGCAGG	2054		
Qy	580	TCCTCTTCTGACTATATCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGAATTTAC	639		
Db	2055	TCCTCTTCTGACTATATCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGAATTTAC	2114		
Qy	640	ATTGTATGGGATGCTTACAGGCTCATGATCTACAGCTGGAAGAAATATCTCTACTGT	699		

Db 2115 ATGTATGGGATGCTCTACAGGCTCATGATCTACAGCTCGGAAGAATAATCTCTACTGT 2174

QY 700 GCTGTTTCATATATGTTGCTCTCAGGTGCGAGTTGGTGAATAATCGGTTTAAAGAGAGTCAA 759

Db 2175 GCTGTTTCATATATGTTGCTCTCAGGTGCGAGTTGGTGAATAATCGGTTTAAAGAGAGTCAA 2234

QY 760 GTATTTCCGCTTGAATACCTAGCCTCTCTAGGTTATGTTGGTTGAGTAGTAGACAACAG 819

Db 2235 GTATTTCCGCTTGAATACCTAGCCTCTCTAGGTTATGTTGGTTGAGTAGTAGACAACAG 2294

QY 820 GGGATCCTGTCAACGAGGGCTTAAATTTGAAGCGGCTTAAATATATAAAATGGCGTCAAT 879

Db 2295 GGGATCCTGTCAACGAGGGCTTAAATTTGAAGCGGCTTAAATATATAAAATGGCGTCAAT 2354

QY 880 AGAATTTGACGATCAGTGAAGGACTCCCAATATCTAGCTTCTCGATGATGATTCATGA 939

Db 2355 AGAATTTGACGATCAGTGAAGGACTCCCAATATCTAGCTTCTCGATGATGATTCATGA 2414

QY 940 CTTAGATCGTGGGATCCACGCGTGTCTTATGAGGATACCTCTCCCTGATGGGATT 999

Db 2415 CTTAGATCGTGGGATCCACGCGTGTCTTATGAGGATACCTCTCCCTGATGGGATT 2474

QY 1000 AATCAGAGTCAAGATATCTTCAGGTTGCTTATGCTGGGGCCCAAGTCACTCTGTGGAT 1059

Db 2475 AATCAGAGTCAAGATATCTTCAGGTTGCTTATGCTGGGGCCCAAGTCACTCTGTGGAT 2534

QY 1060 CTTCTATGATACAGGATACACGGA 1083

Db 2535 CTTCTATGATACAGGATACACGGA 2558

RESULT 13

BC040203

LOCUS Homo sapiens dipeptidylpeptidase 8, mRNA linear PRI 07-OCT-2003

DEFINITION IMAGE:4824813, complete cds.

ACCESSION BC040203

VERSION

KEYWORDS

SOURCE MGC.

ORGANISM Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 4535)

Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Uedin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakeley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzyzanski,M.I., Skalska,U., Smalius,D.E., Schnérch,A., Schein,J.E., Jones,S.J. and Marra,M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932

2 (bases 1 to 4535)

Strausberg,R

AUTHORS

Direct Submission

TITLE

JOURNAL

PUBLISHED

REFERENCE

12477932

2 (bases 1 to 4535)

Strausberg,R

AUTHORS

Direct Submission

TITLE

JOURNAL

Submitted (27-NOV-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

REMARK COMMENT

USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgapbe-r@mail.nih.gov

Tissue Procurement: Miklos Palkovits, M.D., Ph.D.

cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki Toshiyuki and Piero Carninci (NIKEN)

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center

Center code: BCM-HGSC

Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>

Contact: amgebcm.tmc.edu

Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H., Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati, A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAK Plate: 71 Row: k Column: 23.

Location/Qualifiers

source

1. 4535

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="MGC:42570 IMAGE:4824813"

/tissue_type="Testis"

/clone_lib="NIH MGC_97"

/lab_host="DH10B"

/notes="Vector: pBluescript"

gene

1. 4535

/gene="DPP8"

/notes="synonyms: FLJ20283, FLJ14920, MGC26191"

/db_xref="GeneID:54878"

/db_xref="MIM:606819"

1623..4277

/gene="DPP8"

/codon_start=1

/product="DPP8 protein"

/protein_id="AAH40203.1"

/db_xref="GI:26007917"

/db_xref="GeneID:54878"

/db_xref="MIM:606819"

/translation="MAAAMETEQLGVEIFETADCEENIESQDRPKLPFFVVERYSWSQ LKLLADTRKHYGMWAKAPHDPMVKENDPDGPHSDRIYVYLAGSGENRENTLIFYSEI PTFINRAAVLMLSWKPLLDLQATLDYMGYSREELRERKRIGTVGIAGSYDHYQSGG TLFQAGSGIYHKVQGGQGTQQLRENVLVETSCPNRMDPKLCPADPDWIAFHGN DWISNVITREERLTYVHNEIANNEDARGAVTFLQEFDRYSGYVWPKAEIT PGGKILRIYVEENDESEIIVHTSPMLETRADSPRYKPTGTANPKVTFKMEIMI DAEGRIDVDKELIQPPEILFGVEYLARAGTPEGYAWSLIDRSQRLQVLIS PELFIVEDVMERQLIESVPSVTPLIYEETDTWINIHDFHVFQSHBEETEF IFASECTGFRHLXKITSLKESKYKRSSGGLPAPSDFKCPKEEIAITSGEWELGR HGSNCTQDEVRLLVFEQTKDSPLEHLYVYVYVNGEVTRLTDRGYSHSCCISOHCD FTISKYDNKNPHCVSLYKLSSPDDPTCTKTFEATILDSAGPLPDYTPPEIFSPES TTGFTLYGLMKYKPHLOPKKYPTVLFIYGGPOVLNNRFGVKVFRNLATLASGVY VVVIDNRGSHRGLKFECAFKYKMGKGLDQVGLQYLASRYDPLDVRGTHGQYV GGLYSIMALMORSDFRVAIAGAPVTLMWIFVDTGYTRYMGHPDQEQYVLGSAVMO AKKFSPEPNRLLLHGFUDENVFAHTSILLSFLVRAGKPYDLQIFQERHSIRVPS GEHYELHLHLQENLGSRIAALKVI"

CDS

misc_feature

3207..4196

/gene="DPP8"

/note="DAP2; Region: Dipeptidyl aminopeptidases/acylaminoacyl-peptidases [Amino acid transport and metabolism]"

/db_xref="CDD:COG1506"

ORIGIN

Query Match 73.1%; Score 791.2; DB 8; Length 4535;

Best Local Similarity 99.0%; Pred. No. 2.1e-184;

Matches 796; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 280 TGAAGGAGGATCATAGTTGATGAAGTCTGATATTTTGAAGGCCCAAGA 339

|||||

Db	3155	TGAGTCTAATATCCAAAGTTGATGAAGTCAGAGCGTGGTATATTTTGAAGGCACCAAGA	3214
Qy	340	CTCCCTCTTAGAGCATCACCTGTAGTACGTAGTACGTAAATCTCGAGAGGTGACAAAG	399
Db	3215	CTCCCTCTTAGAGCATCACCTGTAGTACGTAGTACGTAAATCTCGAGAGGTGACAAAG	3274
Qy	400	GCTGATGACCGTGGTACTACATCTTCTGCTGATCAGTCAGCAGTGTGATCTTTTAT	459
Db	3275	GCTGATGACCGTGGTACTACATCTTCTGCTGATCAGTCAGCAGTGTGATCTTTTAT	3334
Qy	460	AGTAAGTATAGTACCAAGAAATCCACATCTGTGTCCTTTTCAAGCTATCAAGTCC	519
Db	3335	AGTAAGTATAGTACCAAGAAATCCACATCTGTGTCCTTTTCAAGCTATCAAGTCC	3394
Qy	520	TGAAGTACCCCACTTCCAAAACAAAGGAATTTTGGGCCACCATTTTGGATTACGAGG	579
Db	3395	TGAAGTACCCCACTTCCAAAACAAAGGAATTTTGGGCCACCATTTTGGATTACGAGG	3454
Qy	580	TCTCTTCTGATATATCTTCCAGAAATTTTCTTTTGAAGTACTACTGGATTAC	639
Db	3455	TCTCTTCTGATATATCTTCCAGAAATTTTCTTTTGAAGTACTACTGGATTAC	3514
Qy	640	ATTGATGAGGATGCTCTCAAGCTCATGATCTACAGCTGGAAGAAATATCTACTGT	699
Db	3515	ATTGATGAGGATGCTCTCAAGCTCATGATCTACAGCTGGAAGAAATATCTACTGT	3574
Qy	700	GCTGTTTCATATATGTTGGTCTCAGTGCAGTTCGTTGAATATCGGTTTAAAGGAGTCAA	759
Db	3575	GCTGTTTCATATATGTTGGTCTCAGTGCAGTTCGTTGAATATCGGTTTAAAGGAGTCAA	3634
Qy	760	GTAATTCGTTGATATACCTAGCTCTCTAGGTTATGTTGATGATACACCAAG	819
Db	3635	GTAATTCGTTGATATACCTAGCTCTCTAGGTTATGTTGATGATACACCAAG	3694
Qy	820	GGATCTCTGACCGAGGCTTAAATTTGAAGGCGCTTTAAATATAAATATGAGTCAAT	879
Db	3695	GGATCTCTGACCGAGGCTTAAATTTGAAGGCGCTTTAAATATAAATATGAGTCAAT	3754
Qy	880	AGAAATGACGATCAGTGGAGGACTCCAATATCTAGCTTCTCGATATGATTTCAATGA	939
Db	3755	AGAAATGACGATCAGTGGAGGACTCCAATATCTAGCTTCTCGATATGATTTCAATGA	3814
Qy	940	CTTAGATCGTGGGATCCAGGCTGCTTCTAGGAGGATACCTCTCCCTGATGGCATT	999
Db	3815	CTTAGATCGTGGGATCCAGGCTGCTTCTAGGAGGATACCTCTCCCTGATGGCATT	3874
Qy	1000	AATGACAGGTCAGATATCTTCAAGGTTGCTATTGCTGGGCGCCAGTCACTCTGTGGAT	1059
Db	3875	AATGACAGGTCAGATATCTTCAAGGTTGCTATTGCTGGGCGCCAGTCACTCTGTGGAT	3934
Qy	1060	CTTCTATGATACAGGATACACGGA	1083
Db	3935	CTTCTATGATACAGGATACACGGA	3958
RESULT 14			
AV172659			
LOCUS			
DEFINITION Homo sapiens dipeptidyl peptidase IV-related protein-1 (DPP1)			
ACCESSION AV172659.1 GI:27549549			
VERSION			
KEYWORDS mRNA, complete cds.			
SOURCE Homo sapiens (human)			
ORGANISM			
Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;			
Hominidae; Homo.			
1 (bases 1 to 2656)			
REFERENCE			
AUTHORS Qi, S., Akinsanya, K., Riviere, P. and Junien, J.-L.			
TITLE Novel Serine Protease Genes Related To DPP1V			
JOURNAL Patent: US (WO 0231134)-A 18-APR-2002;			
REFERENCE 2 (bases 1 to 2656)			

AUTHORS	Qi, S., Akinsanya, K., Riviere, P. and Junien, J.-L.
TITLE	Direct Submission
JOURNAL	Submitted (04-NOV-2002) Ferring Research Institute, 3550 General Atomics Ct., San Diego, CA 92121, USA
FEATURES	Location/Qualifiers
source	1..2656
	/organism="Homo sapiens"
	/mol_type="mRNA"
gene	/db_xref="taxon:9606"
	/chTosome="15"
	/map="15q22.1-q22.2"
	1..2656
	/gene="DPP1"
CDS	8..2656
	/gene="DPP1"
	/codon_start=1
	/product="dipeptidyl peptidase IV-related protein-1"
	/protein_id="AA017261.1"
	/db_xref="GI:27549550"
	/translation="MAAMETEQLGVIFETACEBNIESQDPKLEPPFYVRYSWQ
	LKLLADTKYHGMKAKPHDFMKRNDPDGSHDRIDRIYYLAMSGENRENTLFYSEI
	PKTNRAAVLMSWKPLLDLFOATLDYMYSRBEELLRKRKRTGTGIASDYHOGSG
	TFLFAGSGIYHVKGQPGQFTQPLSPNIVETSCPNIRMDPKLCPADWDPAFTHSN
	DWISNIVTREERLLTVHNELANMEDASAGVATFVLQEDRYSQVWCPKAEIT
	PSGGKILILYEENDESEVELIHVTSPLSTRADSRFRPKGTANPKVTFKMSIEMI
	DAEGRIIDVIDKELIQPEILFEVGYIARAGWTPGKYAWMSILLDRFSQTRQLVLIS
	PELFTPVDDVMERQRLIESVPDSVTPLLIYEETDIWINIHDIHFVFFSHHEIEIF
	IFASCKTGFRHLVKITSLKSKYKRSGLPAPDFKCPKIEETAITSGSBEVLGR
	HGSNTQDVEVRLLVYFEGTKDQPLERHLVYVGVETRLDRGYSHSCTISOHCD
	PFISKYKQKPHCVSLKLSBDEDDTCKTKSFATILDSAGPLDYPPTPEIFSPES
	TGFTLYGMLYKPHDLPQKFKYPTVLFIYGGPQVQLVNNRPKGVKFRNLTLASLVY
	VVVIDNRGSGHRGLKPEGKFKYMGQIEIDQVEGQYLASRYDFIDLDVRGHHWSY
	GGYLSFALMORSOIFRVAIAGAPVTLWIFDYGYTERYMGHPDQNEQYILGVSAMQ
	AKFSEPNRLLHLLFLDENVHFANTSIILLFLVRACKPYDLQIYPOEHRHSIRVPES
	GEHYELHLLHQLQENLGSRIAAKVI"

ORIGIN

Query Match		72.9%;	Score 789.6;	DB 8;	Length 2656;
Best Local Similarity		98.9%;	Pred. No. 5.4e-184;		
Matches 795;		Conservative	0;	Mismatches	9;
				Indels	0;
				Gaps	0;
Qy	280	TGAAGAGAGGATCATAGTTGATGAGTCAGAGCGTGGTATATTTTGAAGGCACCAAGA	339		
Db	1534	TGGATCTTAATATATCAAGTTGATGAGTCAGAGCGTGGTATATTTTGAAGGCACCAAGA	1593		
Qy	340	CTCCCTTTAGACATCACCTGTACGTAGTCAGTTTACCTAAATCTCGAGAGGTGACAAG	399		
Db	1594	CTCCCTTTAGACATCACCTGTACGTAGTCAGTTTACCTAAATCTCGAGAGGTGACAAG	1653		
Qy	400	GCTGACTGACCGTGGTACTCATTCTTGTGCATCAGTCAGCAGCTGTGACTCTTTAT	459		
Db	1654	GCTGACTGACCGTGGTACTCATTCTTGTGCATCAGTCAGCAGCTGTGACTCTTTAT	1713		
Qy	460	AAGTAAGTATAGTAAACAGAGNAATCCACACTGTGTCCTTTTACAGCTATCAAGTCC	519		
Db	1714	AAGTAAGTATAGTAAACAGAGNAATCCACACTGTGTCCTTTTACAGCTATCAAGTCC	1773		
Qy	520	TGAAGATGACCCCACTTGCACCAAGAAATTTTGGGCCACCAATTTTGGATTGACGAGG	579		
Db	1774	TGAAGATGACCCCACTTGCACCAAGAAATTTTGGGCCACCAATTTTGGATTGACGAGG	1833		
Qy	580	TCCTCTTCTGACTATATCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGATTAC	639		
Db	1834	TCCTCTTCTGACTATATCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGATTAC	1893		
Qy	640	ATTGTATGGGATGCTCTACAGGCTCATGATCTACAGCTGCAAGAAATATCTTACTGT	699		
Db	1894	ATTGTATGGGATGCTCTACAGGCTCATGATCTACAGCTGCAAGAAATATCTTACTGT	1953		
Qy	700	GCTGTTTCATATATGTTGGTCTCTCAGGTGCAAGTTGGTGAATAATCGGTTTAAAGGAGTCAA	759		
Db	1954	GCTGTTTCATATATGTTGGTCTCTCAGGTGCAAGTTGGTGAATAATCGGTTTAAAGGAGTCAA	2013		

```
QY 760 GTATTTCCGCTTGAATACCCCTAGCCTCTCTAGGTTATGTGGTTGTAGTGATAGACAACAG 819
Db 2014 GTATTTCCGCTTGAATACCCCTAGCCTCTCTAGGTTATGTGGTTGTAGTGATAGACAACAG 2073
QY 820 GGGATCCTGTCAACGAGGCTTAAATTTGAAGCGCCTTAAATATAAATGAGTCAAT 879
Db 2074 GGGATCCTGTCAACGAGGCTTAAATTTGAAGCGCCTTAAATATAAATGAGTCAAT 2133
QY 880 AGAAATTGACGATCAGGTGAAGGACTCCAATATCTAGCTTCTCGATATGATTCATGA 939
Db 2134 AGAAATTGACGATCAGGTGAAGGACTCCAATATCTAGCTTCTCGATATGATTCATGA 2193
QY 940 CTTAGATCGTGGGATCCAGCGTGTCTTATGAGGATACCTCTCCCTGATGGCAAT 999
Db 2194 CTTAGATCGTGGGATCCAGCGTGTCTTATGAGGATACCTCTCCCTGATGGCAAT 2253
QY 1000 AATGCAGAGTCAAGATCTTCAAGGTTGCTATTCCTGGGCCCCAGTCACTCTGTGGAT 1059
Db 2254 AATGCAGAGTCAAGATCTTCAAGGTTGCTATTCCTGGGCCCCAGTCACTCTGTGGAT 2313
QY 1060 CTTCTATGATACAGGATACACGGA 1083
Db 2314 CTTCTATGATACAGGATACACGGA 2337

RESULT 15
AR631278 LOCUS
DEFINITION Sequence 2 from patent US 6844180.
ACCESSION AR631278
VERSION AR631278.1 GI:59770920
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 2671)
AUTHORS Oi,S., Akinsanya,K.O., Riviere,P.J.M. and Junien,J.-L.
TITLE Serine protease genes related to DPPiV
JOURNAL Patent: US 6844180-A 2 18-JAN-2005;
Ferring BV;;
NLX;
FEATURES
source
1..2671
/mol_type="genomic DNA"
ORIGIN
Query Match 72.9%; Score 789.6; DB 6; Length 2671;
Best Local Similarity 98.9%; Pred. No. 5.4e-184;
Matches 795; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 280 TGAAGGAGGATCATAGTGTGATGAAGTCAGAGGCTGGTATATTTTGAAGGCACCAAGA 339
Db 1534 TGGATCTAATATCCNAGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGGCACCAAGA 1593
QY 340 CTCCTCTTTAGAGCATACCTGTAGTACGTAGTACGTAAATCTCTGGAGAGGTGACAAG 399
Db 1594 CTCCTCTTTAGAGCATACCTGTAGTACGTAGTACGTAAATCTCTGGAGAGGTGACAAG 1653
QY 400 GCTGACTGACCGTGGCTACTCAATCTTCTGTCATCAGTCAGCACTGTGACTTCTTTAT 459
Db 1654 GCTGACTGACCGTGGCTACTCAATCTTCTGTCATCAGTCAGCACTGTGACTTCTTTAT 1713
QY 460 AAGTAAGPATAGTAACAGAGAATCCACACTGTGTGCCCTTTACAAGCTATCAAGTCC 519
Db 1714 AAGTAAGPATAGTAACAGAGAATCCACACTGTGTGCCCTTTACAAGCTATCAAGTCC 1773
QY 520 TGAAGATGACCAACTTGCAGAAATTTGGGCGCACATTTTGGATTGAGCAGG 579
Db 1774 TGAAGATGACCAACTTGCAGAAATTTGGGCGCACATTTTGGATTGAGCAGG 1833
QY 580 TCCTCTTCCTGACTATCTCTCCAGAAATTTTCTTTTGAAGTACTACTGGATTAC 639
|||||
```

```
Db 1834 TCCTCTTCCTGACTATATCTCTCCAGAAAATTTTCTTTTGAAGTACTACTGGATTAC 1893
QY 640 ATTGTATGGGATGCTCTPACAAAGCCTCATGATCTACAGCCTGGAAGAAATATCTTACTGT 699
Db 1894 ATTGTATGGGATGCTCTPACAAAGCCTCATGATCTACAGCCTGGAAGAAATATCTTACTGT 1953
QY 700 GCTCTTCATATATGGTGGTCTCCTCAGGTGCAAGTTGGTGAATAATCGATTTAAAGGAGTCAA 759
Db 1954 GCTCTTCATATATGGTGGTCTCCTCAGGTGCAAGTTGGTGAATAATCGATTTAAAGGAGTCAA 2013
QY 760 GTATTTCCGCTTGAATAACCTAGCCTCTCTAGGTTATGTGGTTGTAGTGATAGACAACAG 819
Db 2014 GTATTTCCGCTTGAATAACCTAGCCTCTCTAGGTTATGTGGTTGTAGTGATAGACAACAG 2073
QY 820 GGGATCCTGTCAACGAGGCTTAAATTTGAAGCGCCTTAAATATAAATGAGTCAAT 879
Db 2074 GGGATCCTGTCAACGAGGCTTAAATTTGAAGCGCCTTAAATATAAATGAGTCAAT 2133
QY 880 AGAAATTGACGATCAGGTGAAGGACTCCAATATCTAGCTTCTCGATATGATTCATGA 939
Db 2134 AGAAATTGACGATCAGGTGAAGGACTCCAATATCTAGCTTCTCGATATGATTCATGA 2193
QY 940 CTTAGATCGTGGGATCCAGCGTGTCTTATGAGGATACCTCTCCCTGATGGCAAT 999
Db 2194 CTTAGATCGTGGGATCCAGCGTGTCTTATGAGGATACCTCTCCCTGATGGCAAT 2253
QY 1000 AATGCAGAGTCAAGATCTTCAAGGTTGCTATTCCTGGGCCCCAGTCACTCTGTGGAT 1059
Db 2254 AATGCAGAGTCAAGATCTTCAAGGTTGCTATTCCTGGGCCCCAGTCACTCTGTGGAT 2313
QY 1060 CTTCTATGATACAGGATACACGGA 1083
Db 2314 CTTCTATGATACAGGATACACGGA 2337
```

Search completed: May 4, 2006, 18:52:12
Job time : 5400 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 4, 2006, 02:50:44 ; Search time 609.925 Seconds
(without alignments)

11834.025 Million cell updates/sec

Title: US-10-825-632-8

Perfect score: 1083

Sequence: 1 ggaagaagatgccagatcag.....tatgatacaggatacacgga 1083

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_21.*

- 1: geneseqn1980s.*
- 2: geneseqn1990s.*
- 3: geneseqn2000s.*
- 4: geneseqn2001as.*
- 5: geneseqn2001bs.*
- 6: geneseqn2002as.*
- 7: geneseqn2002bs.*
- 8: geneseqn2003as.*
- 9: geneseqn2003bs.*
- 10: geneseqn2003cs.*
- 11: geneseqn2003ds.*
- 12: geneseqn2004as.*
- 13: geneseqn2004bs.*
- 14: geneseqn2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1083	100.0	1083	4 AAC85697	Aac85697 Nucleotid
2	1055.8	97.5	4309	6 ABK83328	Abk83328 cDNA enco
3	791.2	73.1	1821	6 ABV76411	Abv76411 Dipeptidy
4	791.2	73.1	2083	7 ADR41222	Adr41222 Human CD-
5	791.2	73.1	2643	6 AAH99935	Aah99935 Coding se
6	791.2	73.1	2649	8 ABX12255	Abx12255 cDNA enco
7	791.2	73.1	2649	14 ADV43981	Adv43981 Human psy
8	791.2	73.1	2797	12 ADL13374	Adl13374 Human ste
9	791.2	73.1	2929	10 ACA92425	Aca92425 DNA enco
10	791.2	73.1	3106	6 ABK12892	Abk12892 Human pro
11	791.2	73.1	3120	4 AAC85694	Aac85694 Nucleotid
12	791.2	73.1	3120	6 AAD38956	Aad38956 Human dip
13	791.2	73.1	3143	6 AAH99934	Aah99934 cDNA enco
14	789.6	72.9	2671	6 ABK83322	Abk83322 cDNA enco
15	779.2	71.9	4829	6 ABK83327	Abk83327 cDNA enco
16	608.8	56.2	2251	10 ADE79035	Ade79035 Human pro
17	581.8	53.7	2463	10 ACA92424	Aca92424 DNA enco
18	581.8	53.7	2510	6 AAD23843	Aad23843 Human pro
19	581.2	53.7	2952	10 ACA92421	Aca92421 DNA enco

20	579.6	53.5	2161	4 AAH15009	Aah15009 Human cDN
21	579.2	53.5	2696	12 ADI16386	Adi16386 Human pro
22	497.2	45.9	4685	6 ABK83332	Abk83332 cDNA enco
23	479.2	42.4	4676	6 ABK83331	Abk83331 cDNA enco
24	459	42.4	1197	4 AAC85695	Aac85695 Nucleotid
25	434.2	40.1	1669	4 AAC85696	Aac85696 Nucleotid
26	434.2	40.1	2349	14 ADV43982	Adv43982 Human psy
27	434.2	40.1	2668	6 ABN59775	Abn59775 Novel hum
28	434.2	40.1	2842	6 ABN59774	Abn59774 Novel hum
29	428.2	39.5	4523	6 ABK83325	Abk83325 cDNA enco
30	408.6	37.7	2702	13 ADT04072	Adt04072 Human pro
31	365.4	33.7	502	9 ACH32373	Ach32373 Human end
32	345.8	31.9	2952	6 ABK85090	Abk85090 DNA enco
33	345	31.9	1856	10 AD507574	Ad507574 Novel cod
34	344.2	31.8	2495	6 AAD38957	Aad38957 Human dip
35	344.2	31.8	2617	6 ABK83323	Abk83323 cDNA enco
36	344.2	31.8	2751	6 AAD38311	Aad38311 Murine di
37	344.2	31.8	2766	7 ADR41345	Adr41345 Human CD-
38	344.2	31.8	3024	6 AAD38954	Aad38954 Human dip
39	344.2	31.8	3047	6 ABK69113	Abk69113 DNA enco
40	344.2	31.8	3287	6 AAD38955	Aad38955 Alternati
41	344.2	31.8	3716	6 ABO75955	Abq75955 Human PM
42	344.2	31.8	4147	12 ADO86753	Ado86753 Human tum
43	344.2	31.8	4147	13 AD84273	Ad84273 Human tum
44	344.2	31.8	4147	13 ACN40423	Acn40423 Tumour-as
45	344.2	31.8	4219	6 ABK83335	Abk83335 cDNA enco

ALIGNMENTS

RESULT 1
AAC85697
ID AAC85697 standard; cDNA; 1083 BP.
AC AAC85697;
XX
XX 29-JUN-2001 (first entry)
DT Nucleotide sequence encoding human DPP8 244Glu-341Ile+515Val-776Thr.
DE Human; dipeptidyl aminopeptidase; DPP8; prolyl oligopeptidase;
XX dipeptidyl peptidase; DPPIV; T cell; cleavage; diarrhoea;
KW growth hormone deficiency; glucose level; mucosal regeneration;
KW non-insulin dependent diabetes mellitus; glucose intolerance;
KW immunosuppression; ss.
XX Homo sapiens.
OS
XX
FH Key Location/Qualifiers
FT CDS 2..1081
FT /*tag= a
FT /product= "Human DPP8 244Glu-341Ile+515Val-776Thr"
FT /note= "No stop codon given"

therapeutic uses, and for detecting activated T cells.

Claim 21; Page 75-76; 78pp; English.

The sequences given in AAC85695-97 encode fragments of human dipeptidyl aminopeptidase (DPP8). DPP8 has substrate specificity for H-Ala-Pro-pNA, H-Gly-Pro-pNA and H-Arg-Pro-pNA. Therefore, it is a prolyl oligopeptidase and a dipeptidyl peptidase, because it is capable of hydrolysing the peptide bond C-terminal to Pro in each of these compounds. DPP8 is homologous with human DPPIV. DPP8 is useful for cleaving a substrate, and for detecting an activated T cell which involves measuring the level of DPP8 gene expression in a T cell. The level of DPP8 expression is detected by detecting the amount of DPP8 RNA in the cell. It is also useful for identifying a molecule capable of inhibiting the cleavage of the substrate by DPP8. Molecules identified as inhibiting DPP8 catalytic activity may be useful for treating diarrhoea, growth hormone deficiency, lowering glucose levels in non-insulin dependent diabetes mellitus and other disorders involving glucose intolerance, enhancing mucosal regeneration and as immunosuppressants

Sequence 1083 BP; 312 A; 213 C; 247 G; 311 T; 0 U; 0 Other;

Query Match 100.0%; Score 1083; DB 4; Length 1083;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1083; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	GGAAGAAGATGCCAGATCAGCTGGAGTCGCTACCTTGTCTCCAAAGAAGAAATTGATAG	60
DB	1	GGAAGAAGATGCCAGATCAGCTGGAGTCGCTACCTTGTCTCCAAAGAAGAAATTGATAG	60
QY	61	ATATTTCTGGCTATTGGTGTGTCCTCAAAAGCTGAAACAACTCCCAAGTGGTGTAATAATCT	120
DB	61	ATATTTCTGGCTATTGGTGTGTCCTCAAAAGCTGAAACAACTCCCAAGTGGTGTAATAATCT	120
QY	121	TAGAAATTTCTATAGAAGAAATGATGAATCTGAGGTGGAAATTTATCATGTTACATCCCC	180
DB	121	TAGAAATTTCTATAGAAGAAATGATGAATCTGAGGTGGAAATTTATCATGTTACATCCCC	180
QY	181	TATCTTGGAAACAGAGGGCAGATTCAATTCGGTTATCTCTAAACAGGTACAGCAAAATCC	240
DB	181	TATCTTGGAAACAGAGGGCAGATTCAATTCGGTTATCTCTAAACAGGTACAGCAAAATCC	240
QY	241	TAAAGTCACATTTTAAAGATGTCAGAAATATGATGATGCTGAAGAGGATCATAGTTGA	300
DB	241	TAAAGTCACATTTTAAAGATGTCAGAAATATGATGATGCTGAAGAGGATCATAGTTGA	300
QY	301	TGAAGTCAGAGGCTGTATATTTTGAAGGCCCAAGACTCCCTTTAGAGCATCACT	360
DB	301	TGAAGTCAGAGGCTGTATATTTTGAAGGCCCAAGACTCCCTTTAGAGCATCACT	360
QY	361	GTACGTAGTCAGTTACGTAAATCTCGAGAGGTGACAGGCTGACTGACCGTGTACTC	420
DB	361	GTACGTAGTCAGTTACGTAAATCTCGAGAGGTGACAGGCTGACTGACCGTGTACTC	420
QY	421	ACATTTCTGTGCATCAGTCAGCACTGTGACTCTTTTATAGTAAGTATAGTAAACAGAA	480
DB	421	ACATTTCTGTGCATCAGTCAGCACTGTGACTCTTTTATAGTAAGTATAGTAAACAGAA	480
QY	481	GAATCCACACTGTGTGTCCTTTTACAAGCTATCAAGTCTGGAAGATGACCCAACTTGCAA	540
DB	481	GAATCCACACTGTGTGTCCTTTTACAAGCTATCAAGTCTGGAAGATGACCCAACTTGCAA	540
QY	541	AACAAAGAAATTTTGGGCCACCAATTTTGGATTGAGGATGCTCTCTCTGACTATCTCC	600
DB	541	AACAAAGAAATTTTGGGCCACCAATTTTGGATTGAGGATGCTCTCTCTGACTATCTCC	600
QY	601	TCCAGAAATTTTCTCTTTTCAAAAGTACTCTGATTTACATGTTATGGATGCTCTACAA	660
DB	601	TCCAGAAATTTTCTCTTTTCAAAAGTACTCTGATTTACATGTTATGGATGCTCTACAA	660
QY	661	GCCTCATGATCTACAGCTCGAAAGAAATATCTACTGTGCTGTTTCATATATGTTGCTCC	720
DB	661	GCCTCATGATCTACAGCTCGAAAGAAATATCTACTGTGCTGTTTCATATATGTTGCTCC	720

QY	721	TCAGGTCAGCTTGGTGAATAATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCCCT	780
DB	721	TCAGGTCAGCTTGGTGAATAATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCCCT	780
QY	781	AGCCTCTCTAGGTTATGTGGTTGTAGTGATAGACAAACAGGGGATCCTCTCACCGAGGCT	840
DB	781	AGCCTCTCTAGGTTATGTGGTTGTAGTGATAGACAAACAGGGGATCCTCTCACCGAGGCT	840
QY	841	TAAATTTGAAAGGGCCCTTTTAAATATATAAAATGGGTCAAAATAGAAATTCAGATCAGGTGA	900
DB	841	TAAATTTGAAAGGGCCCTTTTAAATATATAAAATGGGTCAAAATAGAAATTCAGATCAGGTGA	900
QY	901	AGGACTCCAATATCTAGCTTCTCGATATGATTTTCACTTAGATCTGTGGGATCCCA	960
DB	901	AGGACTCCAATATCTAGCTTCTCGATATGATTTTCACTTAGATCTGTGGGATCCCA	960
QY	961	CGGCTGGTCTTATGGAGGATACCTCTCCCTGATGGCATTAATGCAGAGTTCAGATATCTT	1020
DB	961	CGGCTGGTCTTATGGAGGATACCTCTCCCTGATGGCATTAATGCAGAGTTCAGATATCTT	1020
QY	1021	CAGGTTGCTTATGGTGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGATACAC	1080
DB	1021	CAGGTTGCTTATGGTGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGATACAC	1080
QY	1081	GGA 1083	
DB	1081	GGA 1083	
RESULT 2			
ABK83328			
ID	ABK83328	standard; cDNA; 4309 BP.	
XX	AC	ABK83328;	
XX	DT	12-AUG-2002 (first entry)	
XX	DE	cDNA encoding human DPRP-1 splice variant #4.	
XX	KW	Human; serine protease; dipeptidyl peptidase IV-related protein; DPRP;	
XX	KW	DPPIV; infection; human immunodeficiency virus; HIV-1; HIV-2; pain;	
XX	KW	diabetes; infertility; obesity; anorexia; Parkinson's disease; stroke;	
XX	KW	heart failure; hypertension; urinary retention; osteoporosis; cancer;	
XX	KW	ulcer; allergy; cancer; psychotic disorder; neurological disorder;	
XX	KW	dyskinesia; reproductive disorder; inflammatory disorder;	
XX	OS	metabolic disorder; gene; ss.	
XX	XX	Homo sapiens.	
XX	PN	WO200231134-A2.	
XX	PD	18-APR-2002.	
XX	PF	12-OCT-2001; 2001WO-US031874.	
XX	PR	12-OCT-2000; 2000US-0240117P.	
XX	XX	(FERR) FERRING BV.	
XX	PI	Qi S, Akinsanya KO, Riviere PU, Junien J;	
XX	DR	WPI; 2002-444178/47.	
XX	DR	P-PSDB; ABG61597.	
XX	PT	New dipeptidyl peptidase IV-related proteins and nucleic acids encoding	
XX	PT	the proteins, useful for treating e.g. fungal, bacterial, protozoan and	
XX	PT	viral infections, cancers, allergies, neurological disorders, or pain.	
XX	PS	Disclosure; Page 67-68; 113pp; English.	
XX	CC	The present invention relates to the isolation of novel human serine	
XX	CC	proteases referred to as dipeptidyl peptidase IV (DPPIV)-related proteins	

CC (DPRP). The dipeptidyl peptidase IV-related proteins (DPRP) and nucleic
CC acids encoding them are useful for treating infections such as fungal,
CC bacterial, protozoan and viral infections, particularly infections caused
CC by human immunodeficiency virus (HIV-1 or HIV-2), pain, diabetes,
CC precocious puberty, infertility, obesity, anorexia, bulimia, Parkinson's
CC disease, acute heart failure, hypotension, hypertension, urinary
CC retention, osteoporosis, angina pectoris, stroke, ulcers, asthma,
CC allergies, cancers, migraine, vomiting, psychotic and neurological
CC disorders (e.g. anxiety, dementia, or schizophrenia), and dyskinesias.
CC These may also be used in discovering therapeutic agents for the
CC treatment of reproductive, inflammatory and metabolic disorders. ABK83322
CC -ABK83343 encode human DPRP proteins
XX
SQ Sequence 4309 BP; 1304 A; 790 C; 907 G; 1308 T; 0 U; 0 Other;

Query Match 97.5%; Score 1055.8; DB 6; Length 4309;
Best Local Similarity 99.5%; Pred. No. 7.5e-313;
Matches 1080; Conservative 0; Mismatches 2; Indels 3; Gaps 2;

QY 1 GGAAGAAGATCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAAAGAAGATTGTATAG 60
DB 942 GGAAGAAGATCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAAAGAAGATTGTATAG 1001
QY 61 ATATTTCTGGCTATTGGTGGTCCAAAGCTGAACAACTCCACAGTGGTGTAAATTTCT 120
DB 1002 ATATTTCTGGCTATTGGTGGTCCAAAGCTGAACAACTCCACAGTGGTGTAAATTTCT 1061
QY 121 TAGAATTCCTATATGAAGAAATGATGAATCTGAGGTGGAAATTTATTCATGTTACATCCC 180
DB 1062 TAGAATTCCTATATGAAGAAATGATGAATCTGAGGTGGAAATTTATTCATGTTACATCCC 1121
QY 181 TAATGTTGAAACAAGAGGCGAGATTCATTCCTGTTATCTCTAAACAGGTACAGCAATCC 240
DB 1122 TAATGTTGAAACAAGAGGCGAGATTCATTCCTGTTATCTCTAAACAGGTACAGCAATCC 1181
QY 241 TAAGTCACTTTTAAGATGTCAGAAATATGATGATGATGATGATGATGATGATGATGATG 300
DB 1182 TAAGTCACTTTTAAGATGTCAGAAATATGATGATGATGATGATGATGATGATGATGATG 1240
QY 301 TGAAGTCAGAAGCTGGTATATTTTGAAGCCCAAGACTCCCTTTAGAGCATCACCT 360
DB 1241 TGAAGTCAGAAGCTGGTATATTTTGAAGCCCAAGACTCCCTTTAGAGCATCACCT 1300
QY 361 GTAGTAGTCAGTTACGTAATCTCGAGAGGTGACAGGCTGACTGACCGTGGTACTC 420
DB 1301 GTAGTAGTCAGTTACGTAATCTCGAGAGGTGACAGGCTGACTGACCGTGGTACTC 1360
QY 421 ACATTTCTGTCGATCAGTCAGCTGAGCTGACTTCTTTATAGTAAGTATAGTAACCGAA 480
DB 1361 ACATTTCTGTCGATCAGTCAGCTGAGCTGACTTCTTTATAGTAAGTATAGTAACCGAA 1420
QY 481 GAATCCACATGCTGTGCTCCCTTTTACAGCTATCAAGTCTCTGAAGATGACCCAACTTGCA 540
DB 1421 GAATCCACATGCTGTGCTCCCTTTTACAGCTATCAAGTCTCTGAAGATGACCCAACTTGCA 1480
QY 541 AACAAAGGAATTTTGGGCCACCAATTTTGGATTTCAGAGGTCTCTTCTGCTGACTACTCC 600
DB 1481 AACAAAGGAATTTTGGGCCACCAATTTTGGATTTCAGAGGTCTCTTCTGCTGACTACTCC 1540
QY 601 TCCAGAAATTTCTTTTGAAGTACTACTGATTTTACATTTGATGGATGCTCTACAA 660
DB 1541 TCCAGAAATTTCTTTTGAAGTACTACTGATTTTACATTTGATGGATGCTCTACAA 1600
QY 661 GCCTCATGATCTACAGCTGGAAAGAAATATCTACTGCTGCTGCTCATATATGGTG--T 718
DB 1601 GCCTCATGATCTACAGCTGGAAAGAAATATCTACTGCTGCTGCTCATATATGGTG--T 1660
QY 719 CCTCAGGTCAGTTGGTGAATATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACC 778
DB 1661 CCTCAGGTCAGTTGGTGAATATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACC 1720
QY 779 CTAGCCTCTCTAGGTTATGTTGTTAGTGTATGATAGCAACAGGGGATCTCTGTCACCGAGG 838

DB 1721 CTAGCCTCTCTAGGTTATGTTGTTAGTGTATGATAGCAACAGGGGATCTCTGTCACCGAGG 1780
QY 839 CTTAAATTTGAAGCGCCTTTTAAATATAAAATGGGTCAAATAGAAATTCACGATCAGGTG 898
DB 1781 CTTAAATTTGAAGCGCCTTTTAAATATAAAATGGGTCAAATAGAAATTCACGATCAGGTG 1840
QY 899 GAAAGACTCCCAATATCTAGCTTCTCGATATGATTTTCATTGACTTAGATCGTGGGCATC 958
DB 1841 GAAAGACTCCCAATATCTAGCTTCTCGATATGATTTTCATTGACTTAGATCGTGGGCATC 1900
QY 959 CACGCTGGTCTCTATGAGAGATACCTCTCCCTGATGGCAATTAATGAGAGGTACAGATTC 1018
DB 1901 CACGCTGGTCTCTATGAGAGATACCTCTCCCTGATGGCAATTAATGAGAGGTACAGATTC 1960
QY 1019 TTCAGGTTGCTATTTGCTGGGGCCCGAGTCACCTCTCTGATCTTCTATGATACAGATAC 1078
DB 1961 TTCAGGTTGCTATTTGCTGGGGCCCGAGTCACCTCTCTGATCTTCTATGATACAGATAC 2020
QY 1079 ACGGA 1083
DB 2021 ACGGA 2025
RESULT 3
ABV76411
ID ABV76411 standard; cDNA; 1821 BP.
XX
AC ABV76411;
XX
DT 07-FEB-2003 (first entry)
XX
DE Dipeptidyl peptidase IV protein 27.5-encoding cDNA.
XX
KW Dipeptidyl peptidase IV protein 27.5; recombinant production;
KW Stenotrophomonas maltophilia dipeptidyl peptidase IV homologue;
KW gene therapy; diabetes; tumour; cancer; cytostatic; gene; ss.
XX
OS Unidentified.
XX
FH Key Location/Qualifiers
FT CDS 826..1578
FT /tag= a
FT /product= "Dipeptidyl peptidase IV protein 27.5"
XX
PN CN1342768-A.
XX
PD 03-APR-2002.
XX
PF 12-SEP-2000; 2000CN-00125127.
XX
PR 12-SEP-2000; 2000CN-00125127.
XX
PA (BODE-) BODE GENE DEV CO LTD SHANGHAI.
XX
PI Mao Y, Xie Y;
XX
DR WPI; 2002-529809/57.
XX
DR P-PSDB; ABB99949.
XX
PT New dipeptidyl peptidase IV protein 27.5 and encoding polynucleotide,
XX useful for treating diabetes and tumors.
XX
PS Claim 6; Page 24-25 (Disclosure); 32pp; Chinese.
XX
CC The invention relates to dipeptidyl peptidase IV protein 27.5 (ABB99949)
CC and nucleic acids encoding it (ABV76411). The protein has a molecular
CC weight of 27.5 kD and has 40% identity and 58% homology over a 227 amino
CC acid stretch with Stenotrophomonas maltophilia dipeptidyl peptidase IV
CC (GenBank accession number D83263). The invention also relates to a method
CC for the recombinant production of the protein, an antagonist of the
CC protein, and the use of the protein, gene and antagonist in therapeutic
CC applications. Dipeptidyl peptidase IV protein 27.5 can be used in the
CC treatment of a variety of diseases such as diabetes and tumours. The

CC	present sequence represents cDNA encoding dipeptidyl peptidase IV protein	
CC	27.5	
XX	Sequence 1821 BP; 533 A; 374 C; 391 G; 523 T; 0 U; 0 Other;	
SQ	Query Match 73.1%; Score 791.2; DB 6; Length 1821; Best Local Similarity 99.0%; Pred. No. 9e-232; Matches 796; Conservative 0; Mismatches 8; Indels 0; Gaps 0;	
QY	280 TGAAGGAGGATCATAGTTGATGAAGTCAGAAGGCTGGTGTATATTTTGAAGGCACCAAGA 339	
Db	456 TGGATCTAAATATCAAGTTGATGAAGTCAGAAGGCTGGTGTATATTTTGAAGGCACCAAGA 515	
QY	340 CTCGCCCTTTAGAGCATCACCTGTACGTAGTTCAGTTACGTAAATCCCTGGAGAGGTGACAAG 399	
Db	516 CTCGCCCTTTAGAGCATCACCTGTACGTAGTTCAGTTACGTAAATCCCTGGAGAGGTGACAAG 575	
QY	400 GTCGACTGACCGTGGCTACTACATTTCTTGGTGCATCAGTCAGCACTGTGACTTCTTTTAT 459	
Db	576 GTCGACTGACCGTGGCTACTACATTTCTTGGTGCATCAGTCAGCACTGTGACTTCTTTTAT 635	
QY	460 AAGTAAGTATAGTAACACAGAAGAAATCCACACTGTGTGTCCTTTTACAAGCTATCAAGTCC 519	
Db	636 AAGTAAGTATAGTAACACAGAAGAAATCCACACTGTGTGTCCTTTTACAAGCTATCAAGTCC 695	
QY	520 TGAAGATGACCCACTTGCACCAACCAAGGAATTTTGGGCCACCAATTTTGGATTTCAGCAGG 579	
Db	696 TGAAGATGACCCACTTGCACCAACCAAGGAATTTTGGGCCACCAATTTTGGATTTCAGCAGG 755	
QY	580 TCCTCTTCCTGACTATCTCTCCAGAAATTTCTCTTTGAAAGTACTACTGGATTTAC 639	
Db	756 TCCTCTTCCTGACTATCTCTCCAGAAATTTCTCTTTGAAAGTACTACTGGATTTAC 815	
QY	640 ATTGTATGGATGCTCTACAAGCCTCATGATCTACAGCCTGGAAGAAATATCTCTACTGT 699	
Db	816 ATTGTATGGATGCTCTACAAGCCTCATGATCTACAGCCTGGAAGAAATATCTCTACTGT 875	
QY	700 GCTCTTCATATATGGTGGTCCCTCAGGTGCAAGTGTGTGAATAATCGGTTTAAAGAGTCAA 759	
Db	876 GCTCTTCATATATGGTGGTCCCTCAGGTGCAAGTGTGTGAATAATCGGTTTAAAGAGTCAA 935	
QY	760 GTATTTCCGCTTGAATACCCCTAGCCTCTCTAGGTTATGTGTTGTAGTGATAGACAACAG 819	
Db	936 GTATTTCCGCTTGAATACCCCTAGCCTCTCTAGGTTATGTGTTGTAGTGATAGACAACAG 995	
QY	820 GGGATCCTGTCCACCGAGGCTTAAATTTGAAGCGCGCTTTAAATATAAAATGGGTCAAAT 879	
Db	996 GGGATCCTGTCCACCGAGGCTTAAATTTGAAGCGCGCTTTAAATATAAAATGGGTCAAAT 1055	
QY	880 AGAAATTTGACGATCAGGTGGAAGCACTCCAATATCTAGCTTCTCGATATGATTTCAATTGA 939	
Db	1056 AGAAATTTGACGATCAGGTGGAAGCACTCCAATATCTAGCTTCTCGATATGATTTCAATTGA 1115	
QY	940 CTTAGATCGTGTGGGCATCCACGGCTGGTCTATGAGGATACCTCTCCCTGATGGCATT 999	
Db	1116 CTTAGATCGTGTGGGCATCCACGGCTGGTCTATGAGGATACCTCTCCCTGATGGCATT 1175	
QY	1000 AATCAGAGGTTCAGATATCTTCAGGGTTGCTATTGCTGGGGCCCGAGTCATCTCTGTGGAT 1059	
Db	1176 AATCAGAGGTTCAGATATCTTCAGGGTTGCTATTGCTGGGGCCCGAGTCATCTCTGTGGAT 1235	
QY	1060 CTTCTATCATACAGGATACAGGA 1083	
Db	1236 CTTCTATCATACAGGATACAGGA 1259	
RESULT 4		
ADRA4222/c		
ID	ADRA4222 standard; cDNA; 2083 BP.	
XX		
AC	ADRA1222;	
XX		
DT	07-OCT-2004 (first entry)	

XX	Human CD-like molecule HQDES11 cDNA, SEQ ID NO:21.	
DE	Human; CD-like molecule; cluster of differentiation; diagnosis; prevention; immune disorder; immunodeficiency; autoimmune disorder; blood-related disorder; haematological disorder; haemostatic disorder; thrombolytic disorder; hyperproliferative disorder; cancer; tumour; apoptotic disorder; cardiovascular disorder; respiratory disorder; angiogenic disorder; neovascularisation; neurological disorder; endocrine disorder; reproductive system disorder; infectious disease; gastrointestinal disorder; drug screening; tissue regeneration; chemotaxis; gene therapy; antibody therapy; drug targeting; chromosome mapping; forensic analysis; immunophenotyping; cytostatic; haemostatic; tranquiliser; vulnery; anti-inflamatory; nephrotropic; cardiant; antiallergic; anti-HIV; antirheumatic; antiarthritic; antiprosratic; immunosuppressive; vasotropic; nootropic; neuroprotective; antithyroid; thymimetic; gynaecological; virucide; hepatotropic; antibacterial; dermatological; chromosome 15q22; gene; ss.	
OS	Homo sapiens.	
XX	WO200226930-A2.	
PN	04-APR-2002.	
XX	25-SEP-2001; 2001WO-US029838.	
PF	26-SEP-2000; 2000US-0235484P.	
XX	(HUMA-) HUMAN GENOME SCI INC.	
PA	Rosen CA, Birse CE;	
PI	WPI, 2002-405050/43.	
DR	P-PSDB; ADRA41398.	
XX	Novel polynucleotides and polypeptides useful for treating, preventing or ameliorating cardiovascular, renal, neurovascular, and autoimmune disorders.	
PT	Claim 4; SEQ ID NO 21; 1243pp; English.	
PS	The invention relates to 167 novel human CD (cluster of differentiation)-like molecules (ADRA41388-ADRA41563) and to cDNAs encoding them (seqid:11)-	
CC	Sequence 2083 BP; 590 A; 440 C; 418 G; 635 T; 0 U; 0 Other;	
XX	Query Match 73.1%; Score 791.2; DB 7; Length 2083; Best Local Similarity 99.0%; Pred. No. 9.6e-232; Matches 796; Conservative 0; Mismatches 8; Indels 0; Gaps 0;	
SQ	280 TGAAGGAGGATCATAGTTGATGAAGTCAGAAGGCTGGTGTATATTTTGAAGGCACCAAGA 339	
QY	1401 TGGATCTAAATATCCAAGTTGATGAAGTCAGAAGGCTGGTGTATATTTTGAAGGCACCAAGA 1342	
Db	340 CTCGCCCTTTAGAGCATCACCTGTACGTAGTTCAGTTACGTAAATCTGGAGAGGTGACAAG 399	
QY	1341 CTCGCCCTTTAGAGCATCACCTGTACGTAGTTCAGTTACGTAAATCTGGAGAGGTGACAAG 1282	
Db	400 GCTGACTGACCGTGGCTACTCACATTTCTTGGTGCATCAGTCAGCACTGTGACTTCTTTAT 459	
QY	1281 GCTGACTGACCGTGGCTACTCACATTTCTTGGTGCATCAGTCAGCACTGTGACTTCTTTAT 1222	
Db	460 AAGTAAGTATAGTAACACAGAAGAAATTTTGGGCCACCAATTTTGGATTTCAGCAGG 519	
QY	1221 AAGTAAGTATAGTAACACAGAAGAAATTTTGGGCCACCAATTTTGGATTTCAGCAGG 1162	
Db	520 TGAAGATGACCCACTTGCACCAACCAAGGAATTTTGGGCCACCAATTTTGGATTTCAGCAGG 579	
QY	1161 TGAAGATGACCCACTTGCACCAACCAAGGAATTTTGGGCCACCAATTTTGGATTTCAGCAGG 1102	
Db	580 TCCTCTTCCTGACTATCTCTCCAGAAATTTCTCTTTGAAAGTACTACTGGATTTTAC 639	

Db 1101 TCCTCTTCCTGACTATCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGATTAC 1042
Qy 640 ATTGTATGGATGCTCTCAAGCCTCATGATCTACAGCCTGGAAGAATAATCTCTACTGT 699
Db 1041 ATTGTATGGATGCTCTCAAGCCTCATGATCTACAGCCTGGAAGAATAATCTCTACTGT 982
Qy 700 GCTGTTTATATATGGTGGTCTCTCAGTGAGTGGTGAATATCGGTTTAAAGAGTCAA 759
Db 981 GCTGTTTATATATGGTGGTCTCTCAGTGAGTGGTGAATATCGGTTTAAAGAGTCAA 922
Qy 760 GTATTTCCGCTTGAATACCCCTAGCCTCTCTAGGTTATGTTGGTTGAGTAGAGCAACAG 819
Db 921 GTATTTCCGCTTGAATACCCCTAGCCTCTCTAGGTTATGTTGGTTGAGTAGAGCAACAG 862
Qy 820 GGGATCTGTACCGAGGCTTAAATTTGAAGCGCCTTTAAATATAAAATGGGTCAAAT 879
Db 861 GGGATCTGTACCGAGGCTTAAATTTGAAGCGCCTTTAAATATAAAATGGGTCAAAT 802
Qy 880 AGAAATTCAGATCAGTGAAGGACTCCAAATATCTAGCTTCTCATATGATTTCAATGA 939
Db 801 AGAAATTCAGATCAGTGAAGGACTCCAAATATCTAGCTTCTCATATGATTTCAATGA 742
Qy 940 CTTAGATCGTGTGGGATCCACGGCTGGTCTTATGGAGGATACCTCTCCCTGATGGCAAT 999
Db 741 CTTAGATCGTGTGGGATCCACGGCTGGTCTTATGGAGGATACCTCTCCCTGATGGCAAT 682
Qy 1000 AATCAGAGGTGAGATATCTTTCAGGGTTGCTATTGCTGGGGCCCCAGTCACTCTGTGGAT 1059
Db 681 AATCAGAGGTGAGATATCTTTCAGGGTTGCTATTGCTGGGGCCCCAGTCACTCTGTGGAT 622
Qy 1060 CTTCTATGATACAGGATACAGGA 1083
Db 621 CTTCTATGATACAGGATACAGGA 598

RESULT 5

AAH99935
ID AAH99935 standard; cDNA; 2643 BP.

XX AC AAH99935;

XX DT 12-APR-2002 (first entry)

XX DE Coding sequence of 21953 human prollyl oligopeptidase.

XX KW 21953 prollyl oligopeptidase; antibody; proline; endopeptidase; cancer;
KW cardiovascular disease; autoimmune disease; atopic allergy;
KW neuronal disorder; vascular disorder; prostate disorder; Cystostatic;
KW antidiabetic; antiarthritic; antiasthmatic; antiinflammatory;
KW diabetes mellitus; arthritis; multiple sclerosis; asthma;
KW Grave's disease; neuronal disorder; demyelinating disease; ss.

XX OS Homo sapiens.

XX XX

XX PN WO200179473-A2.

XX XX 11-APR-2001; 2001WO-US040483.

XX XX 18-APR-2000; 2000US-0197508P.

XX XX (MILL-) MILLENNIUM PHARM INC.

XX PI Meyers RA, Williamson M;

XX XX WPI; 2002-034353/04.

DR P-PSDB; AAG78415.

XX PT New polypeptides 21953, member of human prollyl oligopeptidase family,
PT useful as diagnostic targets and therapeutic agents for controlling
PT cancer, lymphoma and leukemia.

XX XX

PS Claim 2; Page; 12lpp; English.

XX This invention relates to an isolated 21953 human prollyl oligopeptidase.
CC Which is cytostatic, antidiabetic, antiarthritic, neuroprotective,
CC antithyroid, dermatological, antipsoriatic, antiasthmatic,
CC ophthalmological, antiinflammatory, nootropic, antiparkinsonian,
CC anticonvulsant, gynecological, vasotropic, antianginal, cardiac,
CC antiatherosclerotic, anorectic and metabolic in its action. Uses include
CC gene therapy, expression or activity of 21953 protein modulator, it is
CC useful for identifying a compound which binds to it and can be used in
CC preventing, treating or detecting a cellular proliferative or
CC differentiative disorder. The 21953 molecules can act as novel diagnostic
CC targets and therapeutic agents for controlling disorders associated with
CC the aberrant activity or degradation of peptide hormones e.g., disorders
CC associated with cell differentiation and proliferation such as cancer,
CC immune function, reproductive, neurological and cardiovascular function.
CC The 21953 molecules are thus useful for treating and preventing cellular
CC proliferative and differentiative disorders, haematopoietic neoplastic
CC disorders, immune disorders such as autoimmune diseases, diabetes
CC mellitus, arthritis, multiple sclerosis, asthma, Grave's disease,
CC neuronal disorders, demyelinating diseases, vascular disorders and
CC metabolism or pain disorders. This sequence represents the cDNA encoding
CC sequence of 21953 human prollyl oligopeptidase. This sequence represents
CC the coding sequence of 21953, being the sequence in between the start and
CC the stop codon of the sequence represented in AAH99934

XX SQ Sequence 2643 BP; 800 A; 514 C; 585 G; 744 T; 0 U; 0 Other;

Query Match 73.1%; Score 791.2; DB 6; Length 2643;

Best Local Similarity 99.0%; Pred. No. 1.1e-231;

Matches 796; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 280 TGAAGGAGGATCATAGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGGCACCAAGA 339
Db 1524 TGGATCTAATATCAAGTTCATGAAGTCAGAGGCTGGTATATTTTGAAGGCACCAAGA 1583
Qy 340 CTCCTCTTTAGAGCATCACCTGTACGTAGTCAGTTACGTAAATCTGGAGAGTGACAAG 399
Db 1584 CTCCTCTTTAGAGCATCACCTGTACGTAGTCAGTTACGTAAATCTGGAGAGTGACAAG 1643
Qy 400 GCTGACTGACCGTGGCTACTCACATCTTGTCTCATCAGTCAGCAGCTGTGACTTCTTAT 459
Db 1644 GCTGACTGACCGTGGCTACTCACATCTTGTCTCATCAGTCAGCAGCTGTGACTTCTTAT 1703
Qy 460 AAGTAAGTATAGTAACCCAGAGAAATCCACTGTGTCCCTTTACAAGCTATCAAGTCC 519
Db 1704 AAGTAAGTATAGTAACCCAGAGAAATCCACTGTGTCCCTTTACAAGCTATCAAGTCC 1763
Qy 520 TGAAGATGACCCCACTTGCAAAACCAAGGAATTTTGGGCCACCATTTGGATTGACGAGG 579
Db 1764 TGAAGATGACCCCACTTGCAAAACCAAGGAATTTTGGGCCACCATTTGGATTGACGAGG 1823
Qy 580 TCCTCTCTCTGACTATATCTCTCAGAAATTTTCTCTTTTGAAGTACTACTGGATTATAC 639
Db 1824 TCCTCTCTCTGACTATATCTCTCAGAAATTTTCTCTTTTGAAGTACTACTGGATTATAC 1883
Qy 640 ATTGTATGGATGCTCTTACAAGCCTCATGATCTACAGCCTGGAAGAATAATCTCTACTGT 699
Db 1884 ATTGTATGGATGCTCTTACAAGCCTCATGATCTACAGCCTGGAAGAATAATCTCTACTGT 1943
Qy 700 GCTGTTTATATATGGTGGTCTCTCAGTGAGTGGTGAATATCGGTTTAAAGAGTCAA 759
Db 1944 GCTGTTTATATATGGTGGTCTCTCAGTGAGTGGTGAATATCGGTTTAAAGAGTCAA 2003
Qy 760 GTATTTCCGCTTGAATACCCCTAGCCTCTCTAGGTTATGTTGGTTGAGTAGATAGACAACAG 819
Db 2004 GTATTTCCGCTTGAATACCCCTAGCCTCTCTAGGTTATGTTGGTTGAGTAGATAGACAACAG 2063
Qy 820 GGGATCTGTGACCGAGGCTTAAATTTGAAGCGCCTTTAAATATAAAATCGGTCAAAT 879
Db 2064 GGGATCTGTGACCGAGGCTTAAATTTGAAGCGCCTTTAAATATAAAATCGGTCAAAT 2123
Qy 880 AGAAATTCAGATCAGTGAAGGACTCCAAATATCTAGCTTCTCGATATGATTTCAATGA 939

DB 2124 AGAATTTGACGATCAGGTGGAAGGACTCCATATCTAGCTTCTCGATATGTTTCATGTA 2183
QY 940 CTTAGATCGGTGGGATCCACGGCTGGTCTTATGAGGATACCTCTCCCTGATGGCAAT 999
DB 2184 CTTAGATCGGTGGGATCCACGGCTGGTCTTATGAGGATACCTCTCCCTGATGGCAAT 2243
QY 1000 AATCCAGAGTCAAGATATCTTCAGGTTCTCTATTCCTGGGGCCCCAGFCATCTCTGTGGAT 1059
DB 2244 AATCCAGAGTCAAGATATCTTCAGGTTCTCTATTCCTGGGGCCCCAGFCATCTCTGTGGAT 2303
QY 1060 CTTCTATGATACAGGATACACGGA 1083
DB 2304 CTTCTATGATACAGGATACACGGA 2327

RESULT 6
ID ABX12255 standard; cDNA; 2649 BP.
XX AC ABX12255;
XX DT 19-MAY-2003 (first entry)
XX DE cDNA encoding human serine protease HIPHUM46.
XX KW Human; ss; gene; HIPHUM46; serine protease; gene therapy; osteoarthritis;
KW serine protease activity modulation; dipeptidyl peptidase activity;
KW musculoskeletal disease; Hepatitis B virus infection; myotonic dystrophy;
KW Alzheimer's disease; paraneuronal palsy; Huntington's disease;
KW amyotrophic lateral sclerosis; malabsorption syndrome; lung disease;
KW irritable bowel syndrome; type I diabetes; faecal incontinence; anaemia;
KW haemorrhoid; proctitis; rectal polyp; small bowel tumour; dyslexia;
KW colorectal tumour; ceroid lipofuscinosis; allergic encephalomyelitis;
XX multiple sclerosis; chromosome 15q21-q22.
OS Homo sapiens.
FH Key
FT CDS 1. 2649
FT /*tag= a
FT /product= "HIPHUM46"
FT /note= "Serine protease"
XX
XX GB2374869-A.
XX 30-OCT-2002.
XX 22-JAN-2002; 2002GB-00001404.
XX 23-JAN-2001; 2001GB-00001760.
XX (GLAX) GLAXO GROUP LTD.
XX Edbrooke MR, Lewis AP;
XX WPI; 2003-150703/15.
XX P-PSDB; ABU07720.
XX
XX Identifying modulators of serine protease activity useful for treating
XX musculoskeletal diseases, by contacting cell expressing a novel serine
XX protease polypeptide with a compound and monitoring serine protease
XX activity.
XX
XX Claim 12; Page 22-26; 38pp; English.
XX
XX The invention relates to a method of identifying a substance that
XX modulates serine protease activity, comprising contacting a cell such as
XX a neuronal cell, lung cell, intestinal cell or a cell infected with a
XX virus, expressing a serine protease polypeptide (HIPHUM 46), or its
XX variant having dipeptidyl peptidase activity, or a serine protease
XX isolated from the cell with a test substance and monitoring for serine
XX protease activity. The method is useful for identifying a substance that

CC modulates serine protease activity. A modulator of the serine protease is
CC useful in the manufacture of a medicament for treatment or prophylaxis of
CC a musculoskeletal disease e.g. osteoarthritis, Hepatitis B virus
CC infection, Alzheimer's disease, paraneuronal palsy, myotonic
CC dystrophy, Huntington's disease or amyotrophic lateral sclerosis.
CC Additional disease that may be treated using modulators of the serine
CC protease include malabsorption syndromes, irritable bowel syndrome, lung
CC disease, type I diabetes, faecal incontinence, haemorrhoids, proctitis,
CC rectal polyps, small bowel tumours, colorectal tumours, anaemia,
CC dyslexia, ceroid lipofuscinosis, allergic encephalomyelitis, and multiple
CC sclerosis. The present invention represents cDNA encoding the human serine
CC protease HIPHUM46 which is located on chromosome 15q21-q22
XX
XX Sequence 2649 BP; 803 A; 514 C; 585 G; 747 T; 0 U; 0 Other;
Query Match 73.1%; Score 791.2; DB 8; Length 2649;
Best Local Similarity 99.0%; Pred. No. 1.1e-231;
Matches 796; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 280 TGAAGGAAGGATCATAGTTGATGAGGTGATGAGGTGATATTTTGAAGGCCACCAAGA 339
DB 1527 TGGATCTAATATCAAGTTGATGAGGTGATGAGGTGATATTTTGAAGGCCACCAAGA 1586
QY 340 CTCCTCTTATAGAGCATCCTGTACGTAGTACGTAAATCTCGAGAGGTGACAAG 399
DB 1587 CTCCTCTTATAGAGCATCCTGTACGTAGTACGTAAATCTCGAGAGGTGACAAG 1646
QY 400 GCTGACTGACCGTGGCTACTACATCTTCTGCTGCATCAGTACGACTGTGACTTCTTAT 459
DB 1647 GCTGACTGACCGTGGCTACTACATCTTCTGCTGCATCAGTACGACTGTGACTTCTTAT 1706
QY 460 AAGTAAGTATAGTAACCAAGAAATCCACACTGTGTGCCCTTTTCAAGCTATCAAGTCC 519
DB 1707 AAGTAAGTATAGTAACCAAGAAATCCACACTGTGTGCCCTTTTCAAGCTATCAAGTCC 1766
QY 520 TGAAGATGACCCAACTTGCAAAACAAAGGAATTTTGGGCCCACTTTTGGATTGACGAG 579
DB 1767 TGAAGATGACCCAACTTGCAAAACAAAGGAATTTTGGGCCCACTTTTGGATTGACGAG 1826
QY 580 TCCTCTTCTGACTATACCTCCAGAAATTTTCTCTTTGAAAGTACTACTGATTTAC 639
DB 1827 TCCTCTTCTGACTATACCTCCAGAAATTTTCTCTTTGAAAGTACTACTGATTTAC 1886
QY 640 ATTGTATGGGATGCTCTCAAGGCTCATGATCTACAGGCTGGAAGAAATATCTTACTGT 699
DB 1887 ATTGTATGGGATGCTCTCAAGGCTCATGATCTACAGGCTGGAAGAAATATCTTACTGT 1946
QY 700 GCTGTTTATATATGTTGGTCTCAGTGAGTGTGTAATATCGGTTTAAAGAGTCAA 759
DB 1947 GCTGTTTATATATGTTGGTCTCAGTGAGTGTGTAATATCGGTTTAAAGAGTCAA 2006
QY 760 GTATTTCCGCTTGAATACCTAGCCTCTCTAGGTTATGTGGTTGTAGTAGACACAG 819
DB 2007 GTATTTCCGCTTGAATACCTAGCCTCTCTAGGTTATGTGGTTGTAGTAGACACAG 2066
QY 820 GGGATCCTGTACCGAGGGCTTAAATTTGAAGGCGCTTTAAATATAAATATGAGTCAAT 879
DB 2067 GGGATCCTGTACCGAGGGCTTAAATTTGAAGGCGCTTTAAATATAAATATGAGTCAAT 2126
QY 880 AGAATTTGAGGATCAGTGTGAGGACTCCATATCTAGCTTCTCGATATGATTTATGTA 939
DB 2127 AGAATTTGAGGATCAGTGTGAGGACTCCATATCTAGCTTCTCGATATGATTTATGTA 2186
QY 940 CTTAGATCGTGTGGGATCCAGGCTGGTCTTATGGAGGATACCTCTCCCTGATGGCAAT 999
DB 2187 CTTAGATCGTGTGGGATCCAGGCTGGTCTTATGGAGGATACCTCTCCCTGATGGCAAT 2246
QY 1000 AATGAGAGGTGAGATATCTTTCAGGTTGTCTATTCTCGGGCCCCAGFCATCTGTGGAT 1059
DB 2247 AATGAGAGGTGAGATATCTTTCAGGTTGTCTATTCTCGGGCCCCAGFCATCTGTGGAT 2306
QY 1060 CTTCTATGATACAGGATACACGGA 1083

```
Db 2307 CTCTATGATACAGGATACACGGA 2330
RESULT 7
ADV43981
ID ADV43981 standard; cDNA; 2649 BP.
XX
AC ADV43981;
XX
DT 10-MAR-2005 (first entry)
XX
DE Human psychoneuroendocrinimmune expressed sequence tag SEQ ID NO 1609.
XX
KW microarray; psychoneuroendocrinimmune; chronic fatigue;
KW non-insulin dependent diabetes; allergy; immune disorder; inflammation;
KW cancer; neoplasm; infection; expressed sequence tag; ss.
XX
OS Homo sapiens.
XX
PN WO2004108899-A2.
XX
PD 16-DEC-2004.
XX
PF 04-JUN-2004; 2004WO-US017686.
XX
PR 04-JUN-2003; 2003US-0475915P.
XX
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Nicholson A, Vernon SD;
XX
WPI; 2005-031682/03.
XX
New microarray comprising probes for genes involved in
PT psychoneuroendocrinimmune (PNI) activity, useful in diagnosing a
PT condition associated with PNI activity, e.g., inflammatory or infectious
PT diseases.
XX
Claim 1; SEQ ID NO 1609; 254pp; English.
XX
The invention relates to a new microarray which comprises probes for
CC genes involved in psychoneuroendocrinimmune (PNI) activity. The
CC microarray is useful in diagnosing a condition associated with PNI
CC activity, such as CPS, type-2 diabetes, allergic condition, inflammation,
CC cancer and infection. The present sequence represents a
CC psychoneuroendocrinimmune gene expressed sequence tag. Note the
CC specification mentions SEQ ID NO of up to 3314 but only sequences up to
CC SEQ ID NO 1829 are provided.
XX
SQ Sequence 2649 BP; 804 A; 514 C; 585 G; 746 T; 0 U; 0 Other;
Query Match 73.1%; Score 791.2; DB 14; Length 2649;
Best Local Similarity 99.0%; Pred. No. 1.1e-231;
Matches 796; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 280 TGAAGAGGATCATAGTTCAGTCAAGTCAGAGCGCTGGTATATTTTGAAGCACCACAAAGA 339
DB 1527 TGGATCTAATATCAAGTTGATGAAGTCAGAGCGCTGGTATATTTTGAAGCACCACAAAGA 1586
QY 340 CTCCTCTTAGAGCATCACCTGTACGTAGTTCAGTTACGTTAAATCCTGGAGAGGTGACAAG 399
DB 1587 CTCCTCTTAGAGCATCACCTGTACGTAGTTCAGTTACGTTAAATCCTGGAGAGGTGACAAG 1646
QY 400 GCTGACTGACCGTGGCTACTCACATTCTGCTGCATCAGTCAGTGTGACTTCTTTAT 459
DB 1647 GCTGACTGACCGTGGCTACTCACATTCTGCTGCATCAGTCAGTGTGACTTCTTTAT 1706
QY 460 AAGTAAGTATAGTACCAAGAGAAATCCACACTGTGTGCTCCCTTTTACAAGCTATCAAGTCC 519
DB 1707 AAGTAAGTATAGTACCAAGAGAAATCCACACTGTGTGCTCCCTTTTACAAGCTATCAAGTCC 1766
QY 520 TGAAGATGACCCCACTTGTGCAAAACAAAGGAATTTTGGGCCACCAATTTTGGATTGACGAGG 579
```

```
1767 TGAAGATGACCCCACTTGTGCAAAACAAAGGAATTTTGGGCCACCAATTTTGGATTGACGAGG 1826
580 TCCTCTTCTGACTATACTCTCTCCAGAAATTTCTCTTTTGAAGTACTACTCTGATTTAC 639
1827 TCCTCTTCTGACTATACTCTCTCCAGAAATTTCTCTTTTGAAGTACTACTCTGATTTAC 1886
640 ATTGTATGGGATGCTCTTACAAGCCTCATGATCTACAGCCTGGAAAGAAATATCTACTGT 699
1887 ATTGTATGGGATGCTCTTACAAGCCTCATGATCTACAGCCTGGAAAGAAATATCTACTGT 1946
700 GCTGTTTCAATATATGTTGCTCTCAGGTGCAAGTTCGTTGAATTAATCGGTTTAAAGGAGTCAA 759
1947 GCTGTTTCAATATATGTTGCTCTCAGGTGCAAGTTCGTTGAATTAATCGGTTTAAAGGAGTCAA 2006
760 GTATTTCCGCTTGAATACCTTAGCTCTCTAGGTTATGTGTTGTTAGTAGTATACACACAG 819
2007 GTATTTCCGCTTGAATACCTTAGCTCTCTAGGTTATGTGTTGTTAGTAGTATACACACAG 2066
820 GGGATCCTGTACACCGAGGCTTAAATTTGAAGGCGCTTTAAATATATAAATGGTCAAAAT 879
2067 GGGATCCTGTACACCGAGGCTTAAATTTGAAGGCGCTTTAAATATATAAATGGTCAAAAT 2126
880 AGAAATTGACGATCAGGTGGAAGGACTCCAAATATCTAGCTTCTCGATATGATTTCAATTGA 939
2127 AGAAATTGACGATCAGGTGGAAGGACTCCAAATATCTAGCTTCTCGATATGATTTCAATTGA 2186
940 CTTAGATCGTGTGGGCATCCACGGCTGGTCCCTATGAGGATACCTCTCCCTGATGGCAAT 999
2187 CTTAGATCGTGTGGGCATCCACGGCTGGTCCCTATGAGGATACCTCTCCCTGATGGCAAT 2246
1000 AATGAGAGGTGAGATATCTTCAGGTTGCTATGCTGGGGCCCCAGTCACTCTGTGGAT 1059
2247 AATGAGAGGTGAGATATCTTCAGGTTGCTATGCTGGGGCCCCAGTCACTCTGTGGAT 2306
1060 CTTCTATGATACAGGATACACGGA 1083
2307 CTTCTATGATACAGGATACACGGA 2330
```

```
RESULT 8
ADL13374
ID ADL13374 standard; cDNA; 2797 BP.
XX
AC ADL13374;
XX
DT 06-MAY-2004 (first entry)
XX
DE Human steroid-induced C3A liver cell cDNA #1103.
XX
KW ss; gene; Hepatotropic; Gene therapy; Wilson disease; liver disorder;
KW steroid therapy; cirrhosis; hepatitis; human; C3A liver cell.
XX
OS Homo sapiens.
XX
PN US6673549-B1.
XX
PD 06-JAN-2004.
XX
PF 12-OCT-2001; 2001US-00976594.
XX
PR 12-OCT-2000; 2000US-0240409P.
XX
PA (INCV-) INCVTE CORP.
XX
PI Furness LM, Buchbinder JL;
XX
WPI; 2004-068610/07.
XX
Combination useful for preparing a composition for treating liver
PT disorders associated with steroid therapy, e.g., cirrhosis or hepatitis,
PT comprises cDNAs that are differentially expressed in response to steroid
PT treatment.
XX
```

PS	Claim 1; SEQ ID NO 1103; 141pp; English.	
XX	The invention relates to a combination comprising cDNAs that are	
CC	differentially expressed in response to steroid treatment. Also included	
CC	are the following: a high throughput method for using a cDNA to detect	
CC	differential expression of nucleic acids in a sample; and a high	
CC	throughput method of screening molecules or compounds to identify a	
CC	ligand that specifically binds a cDNA. The sample is from a subject with	
CC	Wilson disease and comparison of a standard defines a stage of that	
CC	disease. The high throughput method of screening molecules or compounds	
CC	to identify a ligand that specifically binds a cDNA comprises: combining	
CC	the combination with molecules or compounds under conditions to allow	
CC	specific binding; and detecting specific binding between each cDNA and at	
CC	least one molecule or compound. The molecules or compounds are regulatory	
CC	proteins. The combination is useful for preparing a composition for	
CC	treating liver disorders associated with steroid therapy, e.g., cirrhosis	
CC	or hepatitis. The present sequence represents a human cDNA which is	
CC	differentially expressed in steroid-induced C3A liver cells. Note: The	
CC	sequence data for this patent did not form part of the printed	
CC	specification but was obtained in electronic format directly from USPTO	
CC	at seqdata.uspto.gov/sequence.html .	
XX		
SQ	Sequence 2797 BP; 866 A; 551 C; 595 G; 785 T; 0 U; 0 Other;	
	Query Match 73.1%; Score 791.2; DB 12; Length 2797;	
	Best Local Similarity 99.0%; Pred. No. 1.1e-231;	
	Matches 796; Conservative 0; Mismatches 8; Indels 0; Gaps 0;	
QY	280 TGAAGGAAGGATCATAGTTGATGAAGTCAGAAGGCTGGTATATTTTGAAGGCACCAAGA 339	
DB	1411 TGGATCTTAATCAAGTTGATGAAGTCAGAAGGCTGGTATATTTTGAAGGCACCAAGA 1470	
QY	340 CTCCTCTTAGAGCATCACCTGTACGTAGTCAGTTACGTAAATCTCTGGAGAGGTGACAAG 399	
DB	1471 CTCCTCTTAGAGCATCACCTGTACGTAGTCAGTTACGTAAATCTCTGGAGAGGTGACAAG 1530	
QY	400 GCTGACTGACCGTGGCTACTCACATCTTCTGCTCATCAGTCAGCACTGTGACTTCTTTAT 459	
DB	1531 GCTGACTGACCGTGGCTACTCACATCTTCTGCTCATCAGTCAGCACTGTGACTTCTTTAT 1590	
QY	460 AAGTAAGTATAGTAACAGAGAAGTAACACACTGTGTCCCTTTTCAAGCTATCAAGTCC 519	
DB	1591 AAGTAAGTATAGTAACAGAGAAGTAACACACTGTGTCCCTTTTCAAGCTATCAAGTCC 1650	
QY	520 TGAAGATGACCCAACTTGCAAAACAAAGGAATTTTCGGCCACCATTTTGGATTCAGCAGG 579	
DB	1651 TGAAGATGACCCAACTTGCAAAACAAAGGAATTTTCGGCCACCATTTTGGATTCAGCAGG 1710	
QY	580 TCCTCTTCTGACTATCTCTCCAGAAATTTCTCTTTTGAAGTACTACTGATTTAC 639	
DB	1711 TCCTCTTCTGACTATCTCTCCAGAAATTTCTCTTTTGAAGTACTACTGATTTAC 1770	
QY	640 ATTGATCGGATGCTCTCAAGCCTCATGATCTACAGCCTGGAAGAATAATCTCTACTGT 699	
DB	1771 ATTGATCGGATGCTCTCAAGCCTCATGATCTACAGCCTGGAAGAATAATCTCTACTGT 1830	
QY	700 GCTGTTTCATATATGTTGGTCTCTCAGGTGCGAGTTGGTGAATAATTCGGTTTAAAGAGTCAA 759	
DB	1831 GCTGTTTCATATATGTTGGTCTCTCAGGTGCGAGTTGGTGAATAATTCGGTTTAAAGAGTCAA 1890	
QY	760 GTATTTCGCTTGAATACCCCTAGCCTCTCTAGGTTATGTTGGTTAGTAGATAGACAACAG 819	
DB	1891 GTATTTCGCTTGAATACCCCTAGCCTCTCTAGGTTATGTTGGTTAGTAGATAGACAACAG 1950	
QY	820 GGGATCTGTCTCAGGAGGCTTAAATTTGAAGCGGCTTTAAATATAAAATGGGTCAAT 879	
DB	1951 GGGATCTGTCTCAGGAGGCTTAAATTTGAAGCGGCTTTAAATATAAAATGGGTCAAT 2010	
QY	880 AGAAATTCACCATCAGGTGGAAGGACTCCCAATATCTAGCTTCTTCGATATGATTTCAATGA 939	
DB	2011 AGAAATTCACCATCAGGTGGAAGGACTCCCAATATCTAGCTTCTTCGATATGATTTCAATGA 2070	
QY	940 CTTAGATCGTGTGGGCATCCACGGCTGGTCTCTATGAGGATACCTCTCCCTGATGGCATT 999	

Db	2071 CTTAGATCGTGTGGGCATCCACGGCTGGTCTCTATGAGGATACCTCTCCCTGATGGCATT 2130	
QY	1000 AATGCAGAGGTTCAGATATCTTCAGGGTTCGTAATGCTGGGGCCCGAGTCACCTCTGTGGAT 1059	
DB	2131 AATGCAGAGGTTCAGATATCTTCAGGGTTCGTAATGCTGGGGCCCGAGTCACCTCTGTGGAT 2190	
QY	1060 CTTCTATGATACAGGATACACGGA 1083	
DB	2191 CTTCTATGATACAGGATACACGGA 2214	
RESULT 9		
ID	ACA92425	
AC	ACA92425 standard; DNA; 2929 BP.	
AC	ACA92425;	
DT	15-JUL-2003 (first entry)	
XX	DNA encoding human PMMM-10.	
DE		
XX	Human; protein modification and maintenance molecule; PMMM; cancer;	
KW	cell proliferation disorder; atherosclerosis; neurological disorder;	
KW	epilepsy; Huntington's disease; stroke; immune disorder; allergy;	
KW	inflammatory disorder; AIDS; developmental disorder; hypothyroidism;	
KW	Cushing's syndrome; gastrointestinal disorder; epithelial disorder;	
KW	infection; cytostatic; antiarteriosclerotic; anticonvulsant; nootropic;	
KW	neuroprotective; cerebroprotective; anti-HIV; antiallergic; vulnary;	
KW	antiinflammatory; thyromimetic; gene; ds.	
OS	Homo sapiens.	
XX		
PN	WO2003031939-A2.	
XX		
PD	17-APR-2003.	
XX		
PF	11-OCT-2002; 2002WO-US032850.	
PR	12-OCT-2001; 2001US-0329689P.	
PR	25-OCT-2001; 2001US-0335703P.	
PR	09-NOV-2001; 2001US-0348887P.	
PR	28-NOV-2001; 2001US-0334145P.	
PR	06-DEC-2001; 2001US-0337451P.	
PR	14-DEC-2001; 2001US-0340584P.	
XX		
PA	(INCY-) INCYTE GENOMICS INC.	
XX		
PI	Ramkumar J, Gorvad AE, Baughn MR, Emerling BM, Yang J, Lee SY;	
PI	Tran UK, Becha SD, Duggan BM, Lee EA, Griffin JA, Li JX;	
PI	Sprague WW, Hafalia AJA, Chawla NK, Lehr-Nason PM, Kable AE, Yue H;	
PI	Marquis JP, Yao MG, Richardson TW, Tang TY, Jin P, Chien D;	
PI	Bhatia U, Burrill JD, Lee S, Blake JJ, Ho A, Zheng W,	
XX		
DR	WPI; 2003-430274/40.	
DR	P-PSDB; ABU92030.	
XX		
PT	New human protein modification and maintenance molecules (PMMM), useful	
PT	for diagnosing, treating and preventing diseases or conditions associated	
PT	with the aberrant PMMM expression e.g. cancer, atherosclerosis, or	
PT	infections.	
XX		
PS	Claim 5; Page 289; 311pp; English.	
XX		
CC	The present invention relates to the isolation of human protein	
CC	modification and maintenance molecules (PMMM), and the polynucleotide	
CC	sequences encoding them. A total of 40 PMMM polypeptides (designated PMMM	
CC	-1 to PMMM-40) are disclosed. The sequences of the invention are useful	
CC	for diagnosing a condition or disease associated with the expression of	
CC	PMMM in a subject, preparing a polyclonal or monoclonal antibody, and	
CC	generating an expression profile of a sample containing the	
CC	polynucleotides. The diseases or conditions associated with decreased	
CC	expression or overexpression of PMMM are cell proliferation disorders	

CC (e.g. cancer, atherosclerosis), neurological disorders (e.g. epilepsy,
CC Huntington's disease, stroke), immune/inflammatory disorders, (e.g. AIDS,
CC allergies), developmental disorders (e.g. hypothyroidism, Cushing's
CC syndrome), gastrointestinal or epithelial disorders, and infections. The
CC PMM polypeptides or their fragments are useful in screening compounds
CC for effectiveness as agonists or antagonists of the polypeptides, or in
CC altering the expression of the target polynucleotide and compounds that
CC specifically bind to, or modulate the activity of the polypeptide.
CC ACA92416-ACA92455 encode the human PMM polypeptides of the invention
XX
SQ Sequence 2929 BP; 864 A; 606 C; 661 G; 798 T; 0 U; 0 Other;
Query Match 73.1%; Score 791.2; DB 10; Length 2929;
Best Local Similarity 99.0%; Pred. No. 1.1e-231;
Matches 796; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 280 TGAAGGAGGATCATAGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGGCCACCAAGA 339
DB 1555 TGGATCTAATATCCAAGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGGCCACCAAGA 1614
QY 340 CTCCTCCCTTTAGAGCATCACCTGTACGTAGTACGTATACGTTAAATCTCGAGAGGTGACAAG 399
DB 1615 CTCCTCCCTTTAGAGCATCACCTGTACGTAGTACGTATACGTTAAATCTCGAGAGGTGACAAG 1674
QY 400 GCTGACTGACCGTGGCTACTCAATCTTCTGCTGATCAGTCAGCACTGTGACTCTTTTAT 459
DB 1675 GCTGACTGACCGTGGCTACTCAATCTTCTGCTGATCAGTCAGCACTGTGACTCTTTTAT 1734
QY 460 AAGTAAGTATAGTACCAAGAGAAATCCACACTGTGTCTCTTTTACAAGCTATCAAGTCC 519
DB 1735 AAGTAAGTATAGTACCAAGAGAAATCCACACTGTGTCTCTTTTACAAGCTATCAAGTCC 1794
QY 520 TGRAGATGACCCACTTGTCAAAACAAAGGAATTTTGGGCCACCATTTTGGATTTCAGCAGG 579
DB 1795 TGAAGATGACCCACTTGTCAAAACAAAGGAATTTTGGGCCACCATTTTGGATTTCAGCAGG 1854
QY 580 TCCTCTTCTGACTATACTCTCTCAGAAATTTTCTTTTGAAGTACTACTGGATTATAC 639
DB 1855 TCCTCTTCTGACTATACTCTCTCAGAAATTTTCTTTTGAAGTACTACTGGATTATAC 1914
QY 640 ATTGTATGGGATGCTCTACAGGCTCATGATCTACAGCTGCTGAAAGAAATATCCTACTGT 699
DB 1915 ATTGTATGGGATGCTCTACAGGCTCATGATCTACAGCTGCTGAAAGAAATATCCTACTGT 1974
QY 700 GCTGTTTCATATATGGTGGTCTCTCAGGTGCGATTGGTGAATATCGTTTAAAGAGTCAA 759
DB 1975 GCTGTTTCATATATGGTGGTCTCTCAGGTGCGATTGGTGAATATCGTTTAAAGAGTCAA 2034
QY 760 GTATTTTCGGTTGAATACCTAGCTCTCTAGTTATGTTGTAGTGTAGTAGACACAG 819
DB 2035 GTATTTTCGGTTGAATACCTAGCTCTCTAGTTATGTTGTAGTGTAGTAGACACAG 2094
QY 820 GGAATCTGTCCACGGGGCTTAAATTTGAAGCGCTTAAATATAAAATGGGTCAAAT 879
DB 2095 GGAATCTGTCCACGGGGCTTAAATTTGAAGCGCTTAAATATAAAATGGGTCAAAT 2154
QY 880 AGAATTTGACCATCAGGTGGAAGACTTCAATATCTAGTCTTCTCGATGATGTTTCATTGA 939
DB 2155 AGAATTTGACCATCAGGTGGAAGACTTCAATATCTAGTCTTCTCGATGATGTTTCATTGA 2214
QY 940 CTTAGATCGTGTGGGCATCCACGGCTGGTCTTATGAGGATACCTCTCCCTGATGGCATT 999
DB 2215 CTTAGATCGTGTGGGCATCCACGGCTGGTCTTATGAGGATACCTCTCCCTGATGGCATT 2274
QY 1000 AATGCAGAGTTCAGATATCTTCAGGGTTCATATGCTGGGGCCCCAGTCACTCTGTGGAT 1059
DB 2275 AATGCAGAGTTCAGATATCTTCAGGGTTCATATGCTGGGGCCCCAGTCACTCTGTGGAT 2334
QY 1060 CTTCTATGATACAGGATACAGGA 1083
DB 2335 CTTCTATGATACAGGATACAGGA 2358

RESULT 10
ABK12892
ID ABK12892 standard; cDNA; 3106 BP.
XX
AC ABK12892;
XX
DT 09-APR-2002 (first entry)
XX
DE Human protease PRTS-9 cDNA sequence.
XX
KW Human; protease; PRTS; gastrointestinal; Crohn's disease; cancer;
KW cardiovascular; atherosclerosis; autoimmune disorder; dermatitis;
KW inflammatory disorder; acquired immunodeficiency syndrome; AIDS;
KW cell proliferative disorder; developmental disorder; epilepsy;
KW Duchenne muscular dystrophy; epithelial disorder; neurological disorder;
KW reproductive disorder; endometriosis; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
CDS 203..2851
FT /*tag= a
ET /product= "Human protease PRTS-9"
XX
XX WO200198468-A2.
XX
XX 27-DEC-2001.
XX
XX 13-JUN-2001; 2001WO-US019178.
XX
XX 16-JUN-2000; 2000US-0212336P.
PR 22-JUN-2000; 2000US-0213955P.
PR 29-JUN-2000; 2000US-0215396P.
PR 07-JUL-2000; 2000US-0216821P.
PR 14-JUL-2000; 2000US-0218946P.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Yue H, Elliott VS, Gandhi AR, Lal P, Au-Young J, Tribouley CM;
PI Deleage AM, Baughn MR, Nguyen DB, Lee EA, Hafalia A, Khan FA;
PI Walla NK, Yao MG, Lu DM, Patterson C, Tang YT, Walsh RT;
PI Azimzai Y, Lu Y, Ramkumar J, Xu Y, Reddy R, Das D, Kearney L;
PI Kallick DA;
XX
DR WPI; 2002-090437/12.
DR P-PSDB; AAU7479.
XX
XX Twenty one human proteases (referred to as PRTS-1 to PRTS-21), useful in
PT the diagnosis, treatment and prevention of gastrointestinal (e.g.
PT gastritis), cardiovascular (e.g. atherosclerosis) and cell proliferative
PT (e.g. cancer) disorders.
XX
XX Claim 5; Page 166-167; 177pp; English.
XX
XX The present invention relates to twenty one new human proteases, referred
CC to as PRTS-1 to PRTS-21. The PRTS polynucleotides and polypeptides of the
CC invention are useful in the diagnosis, treatment and prevention of
CC gastrointestinal e.g. gastritis, esophageal carcinoma and Crohn's
CC disease, cardiovascular, e.g. atherosclerosis, hypertension and myocardial
CC infarction, autoimmune/inflammatory e.g. acquired immunodeficiency
CC syndrome (AIDS), allergies and osteoarthritis, cell proliferative e.g.
CC cancer, developmental e.g. Duchenne and Becker muscular dystrophy,
CC epithelial e.g. dermatitis, neurological e.g. epilepsy and Alzheimer's
CC disease and reproductive e.g. infertility and endometriosis disorders.
CC Numerous other examples of each disorder are given in the specification.
CC The present nucleic acid sequence encodes the human protease PRTS-9
CC protein of the invention
XX
SQ Sequence 3106 BP; 928 A; 633 C; 704 G; 841 T; 0 U; 0 Other;
Query Match 73.1%; Score 791.2; DB 6; Length 3106;
Best Local Similarity 99.0%; Pred. No. 1.2e-231;
Matches 796; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 280 TGAAGGAGGATCATAGTTCATAGTCAGAGGCTGGTATATATTTTGAAGGCACCAAGA 339
Db 1729 TGGATCTAATATCCAGTTCATAGTTCAGAGGCTGGTATATTTTGAAGGCACCAAGA 1788
QY 340 CTCCTCTTTAGAGATCACCCTGTACGTAGTTCAGTTCAGTAAATCTCGAGAGGTGACAAG 399
Db 1789 CTCCTCTTTAGAGATCACCCTGTACGTAGTTCAGTTCAGTAAATCTCGAGAGGTGACAAG 1848
QY 400 GCTGACTGACCGTGGCTACTCAATCTTGTGCTGATCAGTCAGCAGCTGTGACTTCTTTAT 459
Db 1849 GCTGACTGACCGTGGCTACTCAATCTTGTGCTGATCAGTCAGCAGCTGTGACTTCTTTAT 1908
QY 460 AAGTAAGTATAGTAACCAAGGAATCCACACTGTGTGTCCTTTTCAAGCTATCAAGTCC 519
Db 1909 AAGTAAGTATAGTAACCAAGGAATCCACACTGTGTGTCCTTTTCAAGCTATCAAGTCC 1968
QY 520 TGAAGATGACCCCACTTGCAAAAACAAAGGAATTTTGGGCCACCATTTTGGATTTCAGCAGG 579
Db 1969 TGAAGATGACCCCACTTGCAAAAACAAAGGAATTTTGGGCCACCATTTTGGATTTCAGCAGG 2028
QY 580 TCCTCTCTCTGACTATATCTCTCAGAAAATTTTCTTTTGAAGTACTACTGGATTTC 639
Db 2029 TCCTCTCTCTGACTATATCTCTCAGAAAATTTTCTTTTGAAGTACTACTGGATTTC 2088
QY 640 ATTGTATGGATGCTCTACAAAGCCTCATGATCTACAGCCTGGAAGAAATATCTCTACTGT 699
Db 2089 ATTGTATGGATGCTCTACAAAGCCTCATGATCTACAGCCTGGAAGAAATATCTCTACTGT 2148
QY 700 GCTGTTTCATATATGTTGCTCTCAGGTGAGTTGGTGTATATCGTTTAAAGAGTCAA 759
Db 2149 GCTGTTTCATATATGTTGCTCTCAGGTGAGTTGGTGTATATCGTTTAAAGAGTCAA 2208
QY 760 GTATTTCCGTTGAAATACCTAGCCTCTCTAGTTATGTTGTTAGTGTAGTACACACAG 819
Db 2209 GTATTTCCGTTGAAATACCTAGCCTCTCTAGTTATGTTGTTAGTGTAGTACACACAG 2268
QY 820 GGBATCTGTACAGGCGCTTAAATTTGAAGCGCTTTAAATATAAAATGGGTCAAT 879
Db 2269 GGBATCTGTACAGGCGCTTAAATTTGAAGCGCTTTAAATATAAAATGGGTCAAT 2328
QY 880 AGAAATTCAGATCAGGTGAAGGACTCCAAATCTAGCTTCTCGATATGATTTCAATGA 939
Db 2329 AGAAATTCAGATCAGGTGAAGGACTCCAAATCTAGCTTCTCGATATGATTTCAATGA 2388
QY 940 CTTAGATCGTGTGGGCATCCACGGCTGGTCTTATGGAGGATACCTCTCCCTGATGGCAT 999
Db 2389 CTTAGATCGTGTGGGCATCCACGGCTGGTCTTATGGAGGATACCTCTCCCTGATGGCAT 2448
QY 1000 AATCAGAGGTCAGATATCTTCAGGGTTGCTATTGCTGGGGCCCCAGTCACTCTGTGGAT 1059
Db 2449 AATCAGAGGTCAGATATCTTCAGGGTTGCTATTGCTGGGGCCCCAGTCACTCTGTGGAT 2508
QY 1060 CTTCTATGATACAGGATACACGGA 1083
Db 2509 CTTCTATGATACAGGATACACGGA 2532

RESULT 11
AAC85694
ID AAC85694 standard; cDNA; 3120 BP.
XX
AC AAC85694;
XX
DT 29-JUN-2001 (first entry)
XX
DE Nucleotide sequence of human DPP8.
KW Human; dipeptidyl aminopeptidase; DPP8; prolyl oligopeptidase;
KW dipeptidyl peptidase; DPPIV; T cell; cleavage; diarrhoea;
KW growth hormone deficiency; glucose level; mucosal regeneration;
KW non-insulin dependent diabetes mellitus; glucose intolerance;
KW immunosuppression; ss.

XX OS Homo sapiens.
XX Key Location/Qualifiers
FH CDS 214..2862
FT /*tag= a
FT /product= "Human DPP8"
PN WO200119866-A1.
XX
XX 22-MAR-2001.
XX
XX 11-SEP-2000; 2000WO-AU001085.
XX
XX 10-SEP-1999; 99AU-00002762.
PR 18-FEB-2000; 2000AU-00005709.
XX
XX (UNSY) UNIV SYDNEY.
XX
XX Abbott CA, Gorell MD;
PI
DR WPI; 2001-281520/29.
DR P-PSDB; AAB47187.
XX
XX New human dipeptidyl aminopeptidase (DPP8) useful for cleaving
PT substrates, identifying inhibitors of DPP8 catalytic activity which have
PT therapeutic uses, and for detecting activated T cells.
XX
XX Claim 16; Fig 2; 78pp; English.
XX
XX This sequence encodes human dipeptidyl aminopeptidase (DPP8). DPP8 has
CC substrate specificity for H-Gly-Pro-PNA, H-Gly-Pro-PNA and H-Arg-Pro-PNA.
CC Therefore it is a prolyl oligopeptidase and a dipeptidyl peptidase.
CC because it is capable of hydrolyzing the peptide bond C-terminal to, Pro
CC in each of these compounds. DPP8 is homologous with human DPPIV. DPP8 is
CC useful for cleaving a substrate, and for detecting an activated T cell
CC which involves measuring the level of DPP8 gene expression in a T cell.
CC The level of DPP8 expression is detected by detecting the amount of DPP8
CC RNA in the cell. It is also useful for identifying a molecule capable of
CC inhibiting the cleavage of the substrate by DPP8. Molecules identified as
CC inhibiting DPP8 catalytic activity may be useful for treating insulin
CC growth hormone deficiency, lowering glucose levels in non-insulin
CC dependent diabetes mellitus and other disorders involving glucose
CC intolerance, enhancing mucosal regeneration and as immunosuppressants
XX
SQ Sequence 3120 BP; 936 A; 637 C; 706 G; 841 T; 0 U; 0 Other;
Query Match 73.1%; Score 791.2; DB 4; Length 3120;
Best Local Similarity 99.0%; Pred. No. 1.2e-231;
Matches 796; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 280 TGAAGGAGGATCATAGTTCATAGTTCAGAGGCTGGTATATTTTGAAGGCACCAAGA 339
Db 1740 TGGATCTAATATCCAGTTCATAGTTCAGAGGCTGGTATATTTTGAAGGCACCAAGA 1799
QY 340 CTCCTCTTTAGAGATCACCCTGTACGTAGTTCAGTTCAGTAAATCTCGAGAGGTGACAAG 399
Db 1800 CTCCTCTTTAGAGATCACCCTGTACGTAGTTCAGTTCAGTAAATCTCGAGAGGTGACAAG 1859
QY 400 GCTGACTGACCGTGGCTACTCAATCTTGTGCTGATCAGTCAGCAGCTGTGACTTCTTTAT 459
Db 1860 GCTGACTGACCGTGGCTACTCAATCTTGTGCTGATCAGTCAGCAGCTGTGACTTCTTTAT 1919
QY 460 AAGTAAGTATAGTAACCAAGGAATCCACACTGTGTGTCCTTTTCAAGCTATCAAGTCC 519
Db 1920 AAGTAAGTATAGTAACCAAGGAATCCACACTGTGTGTCCTTTTCAAGCTATCAAGTCC 1979
QY 520 TGAAGATGACCCCACTTGCAAAAACAAAGGAATTTTGGGCCACCATTTTGGATTTCAGCAGG 579
Db 1980 TGAAGATGACCCCACTTGCAAAAACAAAGGAATTTTGGGCCACCATTTTGGATTTCAGCAGG 2039
QY 580 TCCTCTCTCTGACTATATCTCTCAGAAAATTTTCTTTTGAAGTACTACTGGATTTC 639

Db 2040 TCCTCTCTGACTATCTCTCCAGAAATTTCTCTTTTGAAGTACTACTGATTTAC 2099
Qy 640 ATTGTATGGATGCTCTACAGCTCATGATCTACAGCTCGAAGAAATATCTACTGT 699
Db 2100 ATTGTATGGATGCTCTCAAGGCTCATGATCTACAGCTCGAAGAAATATCTACTGT 2159
Qy 700 GCTGTTTCATATATGGTGGTCTCAGTGAGTGGTGAATATCGTTTAAAGAGTCAA 759
Db 2160 GCTGTTTCATATATGGTGGTCTCAGTGAGTGGTGAATATCGTTTAAAGAGTCAA 2219
Qy 760 GTATTTCCGCTTCAATACCTAGCTCTCTAGTGTATGGTGTAGTGATAGACAACAG 819
Db 2220 GTATTTCCGCTTCAATACCTAGCTCTCTAGTGTATGGTGTAGTGATAGACAACAG 2279
Qy 820 GGGATCTCTGACCGAGGCTTAAATTTGAAGGCGCTTTAAATATAAAATGGGTCAAT 879
Db 2280 GGGATCTCTGACCGAGGCTTAAATTTGAAGGCGCTTTAAATATAAAATGGGTCAAT 2339
Qy 880 AGAAATTGACGATCAGTGAAGGACTCCAATATCTAGTCTTCGATATGATTCATGA 939
Db 2340 AGAAATTGACGATCAGTGAAGGACTCCAATATCTAGTCTTCGATATGATTCATGA 2399
Qy 940 CTTAGATCGTGGGATCCAGCGTGGTCTTATGAGGATACCTCTCCCTGATGGCATT 999
Db 2400 CTTAGATCGTGGGATCCAGCGTGGTCTTATGAGGATACCTCTCCCTGATGGCATT 2459
Qy 1000 AATGACAGGTGATATCTTCAAGGTTGCTATGCTGGGCGCCAGTCACTCTGTGGAT 1059
Db 2460 AATGACAGGTGATATCTTCAAGGTTGCTATGCTGGGCGCCAGTCACTCTGTGGAT 2519
Qy 1060 CTTCTATGATACAGGATACACGGA 1083
Db 2520 CTTCTATGATACAGGATACACGGA 2543

RESULT 12
AAD38956
ID AAD38956 standard; cDNA; 3120 BP.
XX AAD38956;
XX AC
XX AC
XX AC
XX 23-SEP-2002 (first entry)
XX Human dipeptidyl peptidase 8 (DPP8) cDNA.
XX Human; dipeptidyl peptidase; DPP; neoplasia; type II diabetes; cirrhosis;
KW autoimmunity; human immuno deficiency virus; HIV infection; cytostatic;
KW graft rejection; antidiabetic; antinflammatory; immunosuppressive;
KW antiviral; enzyme; gene; ss.
XX Homo sapiens.
XX Key Location/Qualifiers
FT CDS 214..2862
FT /*tag= a
FT /product= "Human DPP8 protein"
XX WO200234900-A1.
XX 02-MAY-2002.
XX PD
XX PF 29-OCT-2001; 2001WO-AU001388.
XX PR 27-OCT-2000; 2000AU-00001078.
XX PA (UNSY) UNIV SYDNEY.
XX PI Abbott CA, Gorrell MD;
XX WPI; 2002-454646/48.
XX DR P-PSDB; AAE24170.
XX New dipeptidyl peptidase (DPP) peptides, useful for screening inhibitors

PT of DPP catalytic activity, which may be employed to treat e.g. neoplasia,
PT type II diabetes, cirrhosis, autoimmunity, graft rejection and HIV
XX infection.
PS Example; Fig 1; 91pp; English.
XX The present invention relates to dipeptidyl peptidase (DPP) proteins and
CC polynucleotides encoding such proteins. The DPP peptides are useful for
CC screening inhibitors of DPP catalytic activity. The inhibitors are useful
CC for treating neoplasia, type II diabetes, cirrhosis, autoimmunity, graft
CC rejection and HIV (human immuno deficiency virus) infection. The present
CC sequence is human DPP8 cDNA
XX
SQ Sequence 3120 BP; 936 A; 637 C; 706 G; 841 T; 0 U; 0 Other;
Query Match 73.1%; Score 791.2; DB 6; Length 3120;
Best Local Similarity 99.0%; Pred. No. 1.2e-231;
Matches 796; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 280 TGAAGGAAGATCATAGTTCAGAGGCTGGTATATTTTGAAGGCACCAAGA 339
Db 1740 TGGATCTAATATCAAGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGGCACCAAGA 1799
QY 340 CTCCTCTTTAGAGCATCACCTGTACGTAGTCAGTACGTAAATCTTGGAGAGGTGACAAG 399
Db 1800 CTCCTCTTTAGAGCATCACCTGTACGTAGTCAGTACGTAAATCTTGGAGAGGTGACAAG 1859
QY 400 GCTGACTGACCGTGGTCTACTCACATCTTCTGCTGCATCAGTCAGCACTGTCTTTAT 459
Db 1860 GCTGACTGACCGTGGTCTACTCACATCTTCTGCTGCATCAGTCAGCACTGTCTTTAT 1919
QY 460 AAGTAAGTATAGTAACCAAGGAATCCACACTGTGTCTCCCTTTTACAAGCTATCAAGTCC 519
Db 1920 AAGTAAGTATAGTAACCAAGGAATCCACACTGTGTCTCCCTTTTACAAGCTATCAAGTCC 1979
QY 520 TGAAGATGACCCAACTTTGCAAAAACAAAGGAATTTTGGGCCACCATTTGGATTGACGAG 579
Db 1980 TGAAGATGACCCAACTTTGCAAAAACAAAGGAATTTTGGGCCACCATTTGGATTGACGAG 2039
QY 580 TCCTCTCTGACTATATCTCTCAGAAATTTCTCTTTTGAAGTACTACTGGATTAC 639
Db 2040 TCCTCTCTGACTATATCTCTCAGAAATTTTCTCTTTTGAAGTACTACTGGATTAC 2099
QY 640 ATTGTATGGATGCTCTACAGCTCATGATCTACAGCTCGAAGAAATATCTCTACTGT 699
Db 2100 ATTGTATGGATGCTCTACAGCTCATGATCTACAGCTCGAAGAAATATCTCTACTGT 2159
QY 700 GCTGTTTCATATATGGTGGTCTCTCAGGTGCAAGTGGTGAATATCGTTTAAAGAGTCAA 759
Db 2160 GCTGTTTCATATATGGTGGTCTCTCAGGTGCAAGTGGTGAATATCGTTTAAAGAGTCAA 2219
QY 760 GTATTTCCGCTTGAATACCTTAGCTCTCTAGTGTATGGTGTAGTGATAGACAACAG 819
Db 2220 GTATTTCCGCTTGAATACCTTAGCTCTCTAGTGTATGGTGTAGTGATAGACAACAG 2279
QY 820 GGGATCTCTGTCACCGAGGCTTAAATTTGAAGGCGCTTTAAATATAAAATGGGTCAAT 879
Db 2280 GGGATCTCTGTCACCGAGGCTTAAATTTGAAGGCGCTTTAAATATAAAATGGGTCAAT 2339
QY 880 AGAAATTGACGATCAGTGAAGGACTCCAATATCTAGTCTTCGATATGATTCATGA 939
Db 2340 AGAAATTGACGATCAGTGAAGGACTCCAATATCTAGTCTTCGATATGATTCATGA 2399
QY 940 CTTAGATCGTGGGATCCAGCGTGGTCTCTATGAGGATACCTCTCCCTGATGGCATT 999
Db 2400 CTTAGATCGTGGGATCCAGCGTGGTCTCTATGAGGATACCTCTCCCTGATGGCATT 2459
QY 1000 AATGACAGGTGATATCTTCAAGGTTGCTATGCTGGGCGCCAGTCACTCTGTGGAT 1059
Db 2460 AATGACAGGTGATATCTTCAAGGTTGCTATGCTGGGCGCCAGTCACTCTGTGGAT 2519
QY 1060 CTTCTATGATACAGGATACACGGA 1083

KW diabetes; infertility; obesity; anorexia; Parkinson's disease; stroke;
KW heart failure; hypertension; urinary retention; osteoporosis; cancer;
KW ulcer; allergy; cancer; psychotic disorder; neurological disorder;
KW dyskinesia; reproductive disorder; inflammatory disorder;
KW metabolic disorder; gene; ss.
XX Homo sapiens.
XX WO200231134-A2.
XX 18-APR-2002.
XX 12-OCT-2001; 2001WO-US031874.
XX 12-OCT-2000; 2000US-0240117P.
XX (FERR) FERRING BV.
XX Qi S, Akinsanya KO, Riviere PJ, Junien J;
XX WPI; 2002-444178/47.
XX P-PSDB; ABG61591.
XX New dipeptidyl peptidase IV-related proteins and nucleic acids encoding
PT the proteins, useful for treating e.g. fungal, bacterial, protozoan and
PT viral infections, cancers, allergies, neurological disorders, or pain.
XX Claim 1; Page 53-54; 113pp; English.
XX The present invention relates to the isolation of novel human serine
CC proteases referred to as dipeptidyl peptidase IV (DPPIV)-related proteins
CC (DPRP). The dipeptidyl peptidase IV-related proteins (DPRP) and nucleic
CC acids encoding them are useful for treating infections such as fungal,
CC bacterial, protozoan and viral infections, particularly infections caused
CC by human immunodeficiency virus (HIV-1 or HIV-2), pain, diabetes,
CC precocious puberty, infertility, obesity, anorexia, bulimia, Parkinson's
CC disease, acute heart failure, hypotension, hypertension, urinary
CC retention, osteoporosis, angina pectoris, stroke, ulcers, asthma,
CC allergies, cancers, migraine, vomiting, psychotic and neurological
CC disorders (e.g. anxiety, dementia, or schizophrenia), and dyskinesias.
CC These may also be used in discovering therapeutic agents for the
CC treatment of reproductive, inflammatory and metabolic disorders. ABK83322
CC -ABK83343 encode human DPRP proteins
XX SQ Sequence 2671 BP; 805 A; 524 C; 594 G; 748 T; 0 U; 0 Other;
Query Match 72.9%; Score 789.6; DB 6; Length 2671;
Best Local Similarity 98.9%; Pred. No. 3.4e-231;
Matches 795; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 280 TGAAGGAGGATCATAGTTGATGAGTCAGAGGCTGGTATATTTTGAAGGCACCAAGA 339
DB 1534 TGGATCTAATATCCAAAGTTGATGAGTCAGAGGCTGGTATATTTTGAAGGCACCAAGA 1593
QY 340 CTCCTTTAGAGATCACCCTGACGTAGTACGTAAATCTGGAGAGGTGACAAAG 399
DB 1594 CTCCTTTAGAGATCACCCTGACGTAGTACGTAAATCTGGAGAGGTGACAAAG 1653
QY 400 GTCGACTGACCGTGGTACTCACAATCTTGTGTCATCAGTACGACTGTGACTTCTTTAT 459
DB 1654 GTCGACTGACCGTGGTACTCACAATCTTGTGTCATCAGTACGACTGTGACTTCTTTAT 1713
QY 460 AATTAAGTATAGTAAACAGAGAAATCCACACTGTGTGTCCTTTTCAAGCTATCAAGTCC 519
DB 1714 AATTAAGTATAGTAAACAGAGAAATCCACACTGTGTGTCCTTTTCAAGCTATCAAGTCC 1773
QY 520 TGAAGATGACCAATCTGCAACCAAGGAATTTTGGGCCACATTTTGGATTGACAGG 579
DB 1774 TGAAGATGACCAATCTGCAACCAAGGAATTTTGGGCCACATTTTGGATTGACAGG 1833
QY 580 TCCTCTTCCTGACTATCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTTGGATTTTAC 639
DB 1834 TCCTCTTCCTGACTATCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTTGGATTTTAC 1893

QY 640 ATTGTATGGATGCTCTCAAGCCCTCATGATCTACAGCCTCGAAAGAAATATCCTACTGT 699
DB 1894 ATTGTATGGATGCTCTCAAGCCCTCATGATCTACAGCCTCGAAAGAAATATCCTACTGT 1953
QY 700 GCTGTTTCATATATGTTGGTCTCAGGTGAGTGGTGAATATCGGTTTAAAGGAGTCAA 759
DB 1954 GCTGTTTCATATATGTTGGTCTCAGGTGAGTGGTGAATATCGGTTTAAAGGAGTCAA 2013
QY 760 GTATTTCCGCTTGAATACCTAGCCTCTAGGTTATGTTGTTAGTAGTAGAGCAACAG 819
DB 2014 GTATTTCCGCTTGAATACCTAGCCTCTAGGTTATGTTGTTAGTAGTAGAGCAACAG 2073
QY 820 GGGATCCTGTCAACCGAGGCTTAAATTTGAAGGCGCTTTAAATATAAAATGGGTCAAT 879
DB 2074 GGGATCCTGTCAACCGAGGCTTAAATTTGAAGGCGCTTTAAATATAAAATGGGTCAAT 2133
QY 880 AGAAATGACGATCAGGTGAGGACTCAATATCTAGCTTCTCGATATGATTTCAATGA 939
DB 2134 AGAAATGACGATCAGGTGAGGACTCAATATCTAGCTTCTCGATATGATTTCAATGA 2193
QY 940 CTTAGATCGTGTGGGATCCAGGCTGCTTATGAGGATACCTCTCCCTGATGGCAAT 999
DB 2194 CTTAGATCGTGTGGGATCCAGGCTGCTTATGAGGATACCTCTCCCTGATGGCAAT 2253
QY 1000 AATGCAGAGTCAAGATATCTTTCAGGCTTCTTATGCTGGGCCCCCAGTCACTTCTGTGAT 1059
DB 2254 AATGCAGAGTCAAGATATCTTTCAGGCTTCTTATGCTGGGCCCCCAGTCACTTCTGTGAT 2313
QY 1060 CTTCTATGATACAGGATACACGGA 1083
DB 2314 CTTCTATGATACAGGATACACGGA 2337
RESULT 15
ABK83327
ID ABK83327 standard; cDNA; 4829 BP.
XX AC ABK83327;
XX 12-AUG-2002 (first entry)
XX cDNA encoding human DPRP-1 splice variant #3.
XX Human; serine protease; dipeptidyl peptidase IV-related protein; DPRP;
KW DPPIV; infection; human immunodeficiency virus; HIV-1; HIV-2; pain;
KW diabetes; infertility; obesity; anorexia; Parkinson's disease; stroke;
KW heart failure; hypertension; urinary retention; osteoporosis; cancer;
KW ulcer; allergy; cancer; psychotic disorder; neurological disorder;
KW dyskinesia; reproductive disorder; inflammatory disorder;
KW metabolic disorder; gene; ss.
XX Homo sapiens.
XX WO200231134-A2.
XX 18-APR-2002.
XX 12-OCT-2001; 2001WO-US031874.
XX 12-OCT-2000; 2000US-0240117P.
XX (FERR) FERRING BV.
XX Qi S, Akinsanya KO, Riviere PJ, Junien J;
XX WPI; 2002-444178/47.
XX P-PSDB; ABG61596.
XX New dipeptidyl peptidase IV-related proteins and nucleic acids encoding
PT the proteins, useful for treating e.g. fungal, bacterial, protozoan and
PT viral infections, cancers, allergies, neurological disorders, or pain.
XX

PS Disclosure; Page 65-66; 113pp; English.

XX The present invention relates to the isolation of novel human serine
CC proteases referred to as dipeptidyl peptidase IV (DPPIV)-related proteins
CC (DPRP). The dipeptidyl peptidase IV-related proteins (DPRP) and nucleic
CC acids encoding them are useful for treating infections such as fungal,
CC bacterial, protozoan and viral infections, particularly infections caused
CC by human immunodeficiency virus (HIV-1 or HIV-2), pain, diabetes,
CC precocious puberty, infertility, obesity, anorexia, bulimia, Parkinson's
CC disease, acute heart failure, hypotension, hypertension, urinary
CC retention, osteoporosis, angina pectoris, stroke, ulcers, asthma,
CC allergies, cancers, migraine, vomiting, psychotic and neurological
CC disorders (e.g. anxiety, dementia, or schizophrenia), and dyskinesias.
CC These may also be used in discovering therapeutic agents for the
CC treatment of reproductively inflammatory and metabolic disorders. ABK83322
CC -ABK83343 encode human DPRP proteins
XX

SQ Sequence 4829 BP; 1466 A; 886 C; 1017 G; 1460 T; 0 U; 0 Other;

Query Match 71.9%; Score 779.2; DB 6; Length 4829;
Best Local Similarity 99.8%; Pred. No. 7.2e-228;
Matches 796; Conservative 0; Mismatches 8; Indels 2; Gaps 1;
Qy 280 TGAAGGAAGGATCATAGTTGATGAAGTCAAGAGGCTGGTATATTTTGAAGGCACCAAGA 339
Db |||||
Qy 1740 TGGATCTAAATCAAGTTGATGAAGTCAAGAGGCTGGTATATTTTGAAGGCACCAAGA 1799
Db |||||
Qy 340 CTCCTCTTTAGAGCATCACTGTPACGTAGTCAGTTACGTAAATCCTGGAGAGGTGACAAG 399
Db |||||
Qy 1800 CTCCTCTTTAGAGCATCACTGTPACGTAGTCAGTTACGTAAATCCTGGAGAGGTGACAAG 1859
Db |||||
Qy 400 GCTGACTGACCGTGGCTACTCACATTTCTTGTGTCATCAGTCAGCACTGTGACTTCTTTAT 459
Db |||||
Qy 1860 GCTGACTGACCGTGGCTACTCACATTTCTTGTGTCATCAGTCAGCACTGTGACTTCTTTAT 1919
Db |||||
Qy 460 AAGTAAGTATAGTAACACAGAAGATCCACACTGTGTCCCTTTTACAAGTATCAAGTCC 519
Db |||||
Qy 1920 AAGTAAGTATAGTAACACAGAAGATCCACACTGTGTCCCTTTTACAAGTATCAAGTCC 1979
Db |||||
Qy 520 TGAAGATGACCCAACTTGCACCAAGGAATTTTGGCCACCAATTTGGATTTCAGCAGG 579
Db |||||
Qy 1980 TGAAGATGACCCAACTTGCACCAAGGAATTTTGGCCACCAATTTGGATTTCAGCAGG 2039
Db |||||
Qy 580 TCCTCTTCTGACTACTCTCCAGAAATTTCTCTTTTGAAGTACTACTGGATTTC 639
Db |||||
Qy 2040 TCCTCTTCTGACTACTCTCCAGAAATTTCTCTTTTGAAGTACTACTGGATTTC 2099
Db |||||
Qy 640 ATTGTATGGATGCTCTACAGCCTCATGATCTACGCTTGGAAAGAAATATCCTACTGT 699
Db |||||
Qy 2100 ATTGTATGGATGCTCTACAGCCTCATGATCTACGCTTGGAAAGAAATATCCTACTGT 2159
Db |||||
Qy 700 GCTGTTTCATATATGGTGG--TCCTCAGGTGAGTTGGTGAATATATCGTTTAAAGGAGTC 757
Db |||||
Qy 2160 GCTGTTTCATATATGGTGGTCTCCTCAGGTGAGTTGGTGAATATATCGTTTAAAGGAGTC 2219
Db |||||
Qy 758 AGTATTTCCGTTGGAATACCTAGCCTCTCTAGGTTATGTGGTTGTAGTGATAGACAAC 817
Db |||||
Qy 2220 AGTATTTCCGTTGGAATACCTAGCCTCTCTAGGTTATGTGGTTGTAGTGATAGACAAC 2279
Db |||||
Qy 818 AGGGATCTCTGTACCGAGGGCTTAAATTTGAAGCGCCTTTAAATATAAATGGGTCAA 877
Db |||||
Qy 2280 AGGGATCTCTGTACCGAGGGCTTAAATTTGAAGCGCCTTTAAATATAAATGGGTCAA 2339
Db |||||
Qy 878 ATAGAAATTCAGATCAGGTGAGGACTCCAAATATCTAGCTTCGTGATATGATTTTCATT 937
Db |||||
Qy 2340 ATAGAAATTCAGATCAGGTGAGGACTCCAAATATCTAGCTTCGTGATATGATTTTCATT 2399
Db |||||
Qy 938 GACTTAGATCGTGTGGGATCCACGGCTGGTCCCTATGGAGGATACCTCTCCCTGATGGCA 997
Db |||||
Qy 2400 GACTTAGATCGTGTGGGATCCACGGCTGGTCCCTATGGAGGATACCTCTCCCTGATGGCA 2459
Db |||||
Qy 998 TTAATGCAGAGTCAGATATCTTCAGGGTTGCTATTCTGGGGCCCGCAGTCACTCTGTGG 1057
Db |||||

Db 2460 TTAATGCAGAGTCAGATATCTTCAGGGTTGCTATTCTGGGGCCCGCAGTCACTCTGTGG 2519
Qy 1058 ATCTTCTATGATACAGGATACACCGA 1083
Db |||||
Db 2520 ATCTTCTATGATACAGGATACACCGA 2545
Search completed: May 4, 2006, 09:09:39
Job time : 611.925 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 4, 2006, 02:59:51 ; Search time 4056.93 Seconds
(without alignments)
12489.839 Million cell updates/sec

Title: US-10-825-632-8
Perfect score: 1083
Sequence: 1 ggaagaagatgccagatcag.....tatgatacagatacacgga 1083

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: gb_est1.*
2: gb_est2.*
3: gb_est3.*
4: gb_est4.*
5: gb_est5.*
6: gb_est6.*
7: gb_est7.*
8: gb_est8.*
9: gb_est9.*
10: gb_est10.*
11: gb_est11.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	794.6	73.4	2649	10 AY411615	AY411615 Homo sapi
2	789.8	72.9	2649	10 AY411616	AY411616 Pan trogl
3	789.2	72.9	1042	3 BM557438	BM557438 AGENCOURT
4	727.8	67.2	910	5 BQ675260	BQ675260 AGENCOURT
5	720.6	66.5	957	5 BQ675006	BQ675006 AGENCOURT
6	689	63.6	1265	4 AF176779	AF176779 Homo sapi
7	670.6	61.9	2634	10 AY411617	AY411617 Mus muscu
8	669.6	61.8	3143	4 AK016546	AK016546 Mus muscu
9	669.6	61.8	5517	4 AK029788	AK029788 Mus muscu
10	665.2	61.4	968	5 BQ671635	BQ671635 AGENCOURT
11	662.2	61.1	1041	2 BI084090	BI084090 602869453
12	659.8	60.9	726	7 CN427231	CN427231 170004554
13	642	59.3	1292	4 AF175225	AF175225 Homo sapi
14	619.8	57.2	674	2 BE888665	BE888665 601513061
15	616	56.9	639	2 BG165205	BG165205 602344074
16	607.8	56.1	658	2 BE388695	BE388695 601286345
17	606.6	56.0	753	1 AL040398	AL040398 DKFZp434A
18	596.4	55.1	746	5 BY751026	BY751026 BY751026
19	591.2	54.6	1278	4 AF173382	AF173382 Homo sapi
20	590.6	54.5	766	6 CD803232	CD803232 UI-M-GVO-
21	579.6	53.5	748	6 CD803197	CD803197 UI-M-GVO-
22	578.6	53.4	684	6 CF533690	CF533690 UI-M-FY0-

23	575.2	53.1	729	6	CF726752	CF726752 UI-M-HB0-
24	570.6	52.7	1090	3	BM910838	BM910838 AGENCOURT
25	568.2	52.5	671	8	CX787143	CX787143 HESC3 81
26	565.2	52.2	1013	7	CN647232	CN647232 ILLUMIGEN
27	565.2	52.2	1027	7	CN646612	CN646612 ILLUMIGEN
28	562.2	51.9	683	7	CN530337	CN530337 UI-M-H00-
29	543.8	50.2	853	2	BF694553	BF694553 602081767
30	539.2	49.8	703	6	CF745813	CF745813 UI-M-GVO-
31	535.4	48.4	789	2	BG709118	BG709118 602675382
32	521.4	48.1	695	2	BG715866	BG715866 602675282
33	493.8	45.6	788	2	BG163397	BG163397 602338360
34	492.8	45.5	643	7	CN529315	CN529315 UI-M-H00-
35	486.6	44.9	853	2	BI223892	BI223892 602941035
36	480.8	44.4	897	5	BUI91638	BUI91638 AGENCOURT
37	470.2	43.4	831	2	BG974324	BG974324 602844075
38	454.4	42.0	1082	5	BU239476	BU239476 603322338
39	450.4	41.6	705	7	CO429380	CO429380 UI-M-HW0-
40	434.2	40.1	2292	4	CR609512	CR609512 full-length
41	426.4	39.4	663	7	CN535698	CN535698 UI-M-HS0-
42	409.8	37.8	837	8	CX501844	CX501844 JGI_XZG39
43	409.4	37.8	479	2	BI135177	BI135177 UI-M-BH3-
44	397.4	36.7	738	7	CN427234	CN427234 170006000
45	395.6	36.5	875	8	DN068337	DN068337 JGI_CABD4

ALIGNMENTS

RESULT 1	AY411615	Homo sapiens DPP8 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.	2649 bp	DNA	linear	GSS 16-DEC-2003
AY411615	LOCUS	AY411615	GI:39767583			
DEFINITION	AY411615.1	GI:39767583				
ACCESSION	AY411615					
VERSION	AY411615.1					
KEYWORDS	GSS.					
SOURCE	Homo sapiens (human)					
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euthera; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.					
REFERENCE	1 (bases 1 to 2649)					
AUTHORS	Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarawal, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Ferreria, S., Wang, G., Zheng, X.H., White, T.J., Shinsky, J.J., Adams, M.D. and Cargill, M.					
TITLE	Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios					
JOURNAL	Science 302 (5652), 1960-1963 (2003)					
PUBMED	14671302					
REFERENCE	2 (bases 1 to 2649)					
AUTHORS	Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarawal, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Ferreria, S., Wang, G., Zheng, X.H., White, T.J., Shinsky, J.J., Adams, M.D. and Cargill, M.					
TITLE	Direct Submision					
JOURNAL	Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA					
COMMENT	This sequence was made by sequencing genomic exons and ordering them based on alignment.					
FEATURES	Location/Qualifiers					
source	1..2649					
	/organism="Homo sapiens"					
gene	/mol_type="genomic DNA"					
	/db_xref="taxon:9606"					
	<1..>2649					
	/gene="DPP8"					
	/locus_tag="HCM4283"					
ORIGIN						
Query Match	73.4%;	Score	794.6;	DB	10;	Length 2649;
Best Local Similarity	90.0%;	Pred. No. le-214;				
Matches	800;	Conservative	0;	Mismatches	89;	Indels 0; Gaps 0;

```
QY 195 GGAGGGCAGATTGATTCCGTTATCTCTAAACAGGTACAGCAAACTCTAAAGTCACATTTTA 254
Db 1442 GCTGCCCTGCTCCANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1501
QY 255 AGATGTCAGAAATAATGATTGCTGCTGAAGGAAGGATCATAGTTGATGAAGTCAGAAGGC 314
Db 1502 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1561
QY 315 TGGTATATTTTGAAGGCACCAAGACATCCCTTTAGAGCATCACCTGTACGTAGTCAGTT 374
Db 1562 TGGTATATTTTGAAGGCACCAAGACATCCCTTTAGAGCATCACCTGTACGTAGTCAGTT 1621
QY 375 ACGTAAATCTCTGGAGAGGTGACAAAGGCTGACTGACCGTGGCTACTCACATCTTCTGTGCA 434
Db 1622 ACGTAAATCTCTGGAGAGGTGACAAAGGCTGACTGACCGTGGCTACTCACATCTTCTGTGCA 1681
QY 435 TCAGTCAAGCACTGTGACTCTTTTATAAGTAAGTATAGTAACAGGAAGATCCACACTGTG 494
Db 1682 TCAGTCAAGCACTGTGACTCTTTTATAAGTAAGTATAGTAACAGGAAGATCCACACTGTG 1741
QY 495 TGTCCCTTTTACAAGCTATCAAGTCTCTGAAGATGACCCCACTTGTCAAAACAAGGAATTTT 554
Db 1742 TGTCCCTTTTACAAGCTATCAAGTCTCTGAAGATGACCCCACTTGTCAAAACAAGGAATTTT 1801
QY 555 GGGCCACCATTTTGGATTTCAGCAGGTCCTCTTCTGACTATATCTCTCCAGAAATTTTCT 614
Db 1802 GGGCCACCATTTTGGATTTCAGCAGGTCCTCTTCTGACTATATCTCTCCAGAAATTTTCT 1861
QY 615 CTTTGTGAAGTACTGATGATTTACATTTGATGGATGCTCTACAAAGCTCATGATCTAC 674
Db 1862 CTTTGTGAAGTACTGATGATTTACATTTGATGGATGCTCTACAAAGCTCATGATCTAC 1921
QY 675 AGCTGGAAGAAATATCTTACTGCTGTTCATATATGTTGGTCTCTCAGGTGCAAGTTGG 734
Db 1922 AGCTGGAAGAAATATCTTACTGCTGTTCATATATGTTGGTCTCTCAGGTGCAAGTTGG 1981
QY 735 TGAATAATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCTACGCTCTCTAGGTT 794
Db 1982 TGAATAATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCTACGCTCTCTAGGTT 2041
QY 795 ATGTGGTTGTAGTAGACAAACAGGGGATCTCTGTACCGAGGGCTTAAATTTGAAGGCG 854
Db 2042 ATGTGGTTGTAGTAGACAAACAGGGGATCTCTGTACCGAGGGCTTAAATTTGAAGGCG 2101
QY 855 CTTTAAATATATAAATGGGTCAAATAGAAATTGACGATCAGGTGGAAGGACTCCAATATC 914
Db 2102 CTTTAAATATATAAATGGGTCAAATAGAAATTGACGATCAGGTGGAAGGACTCCAATATC 2161
QY 915 TAGCTTCTCGATATGATTTCAATGACTTAGATCGTGTGGGATCCAGCGCTGTCTTATG 974
Db 2162 TAGCTTCTCGATATGATTTCAATGACTTAGATCGTGTGGGATCCAGCGCTGTCTTATG 2221
QY 975 GAGGATACCTCTCCCTGATGGCATTAAATCGAGAGTTCAGATATCTTCAGGTTGCTATTG 1034
Db 2222 GAGGATACCTCTCCCTGATGGCATTAAATCGAGAGTTCAGATATCTTCAGGTTGCTATTG 2281
QY 1035 CTGGGGCCCGAGTCACTGTGGATCTTCTATGATACAGGATACACGGA 1083
Db 2282 CTGGGGCCCGAGTCACTGTGGATCTTCTATGATACAGGATACACGGA 2330

RESULT 2
LOCUS AY411616 2649 bp DNA linear GSS 16-DEC-2003
DEFINITION Pan troglodytes DPP8 gene, VIRUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY411616
VERSION AY411616.1 GI:39767584
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Pan.
1 (bases 1 to 2649)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,P., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
Science 302 (5652), 1960-1963 (2003)
14671302
2 (bases 1 to 2649)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,P., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering
them based on alignment
Location/Qualifiers
source 1..2649
organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
<!.>2649
/gene="DPP8"
/locus_tag="HCM4283"

Query Match 72.9%; Score 789.8; DB 10; Length 2649;
Best Local Similarity 89.7%; Pred. No. 2.4e-213;
Matches 797; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

QY 195 GGAGGGCAGATTTCATCCGTTATCTCTAAACAGGTACAGCAAACTCTAAAGTCACATTTTA 254
Db 1442 GCTGCCCTGCTCCAAANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1501
QY 255 AGATGTCAGAAATAATGATTGCTGCTGAAGGAAGGATCATAGTTGATGAAGTCAGAAGGC 314
Db 1502 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1561
QY 315 TGGTATATTTTGAAGGCACCAAGACATCCCTTTAGAGCATCACCTGTACGTAGTCAGTT 374
Db 1562 TGGTATATTTTGAAGGCACCAAGACATCCCTTTAGAGCATCACCTGTACGTAGTCAGTT 1621
QY 375 ACGTAAATCTCTGGAGAGGTGACAAAGGCTGACTGACCGTGGCTACTCACATCTTCTGTGCA 434
Db 1622 ACGTAAATCTCTGGAGAGGTGACAAAGGCTGACTGACCGTGGCTACTCACATCTTCTGTGCA 1681
QY 435 TCAGTCAAGCACTGTGACTCTTTTATAAGTAAGTATAGTAACAGGAAGATCCACACTGTG 494
Db 1682 TCAGTCAAGCACTGTGACTCTTTTATAAGTAAGTATAGTAACAGGAAGATCCACACTGTG 1741
QY 495 TGTCCCTTTTACAAGCTATCAAGTCTCTGAAGATGACCCCACTTGTCAAAACAAGGAATTTT 554
Db 1742 TGTCCCTTTTACAAGCTATCAAGTCTCTGAAGATGACCCCACTTGTCAAAACAAGGAATTTT 1801
QY 555 GGGCCACCATTTTGGATTTCAGCAGGTCCTCTTCTGACTATATCTCTCCAGAAATTTTCT 614
Db 1802 GGGCCACCATTTTGGATTTCAGCAGGTCCTCTTCTGACTATATCTCTCCAGAAATTTTCT 1861
QY 615 CTTTGTGAAGTACTGATGATTTACATTTGATGGATGCTCTACAAAGCTCATGATCTAC 674
Db 1862 CTTTGTGAAGTACTGATGATTTACATTTGATGGATGCTCTACAAAGCTCATGATCTAC 1921
QY 675 AGCTGGAAGAAATATCTTACTGCTGTTCATATATGTTGGTCTCTCAGGTGCAAGTTGG 734
Db 1922 AGCTGGAAGAAATATCTTACTGCTGTTCATATATGTTGGTCTCTCAGGTGCAAGTTGG 1981
QY 735 TGAATAATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCTACGCTCTCTAGGTT 794
```

Db 1982 TGAATAAATCGGTTTAAAGGAGCTCAAGTAATTTCCGCTTGAATACCTCTAGCTCTCTAGGTT 2041

QY 795 ATGTGTTGTAGTATGATACACAACAGGGGATCCCTGTCCACGAGGCTTAAATTTGAAGCG 854

Db 2042 ATGTGTTGTAGTATGATACACAACAGGGGATCCCTGTCCACGAGGCTTAAATTTGAAGCG 2101

QY 855 CTTTAAATATAAAATGGTCAAAATAGAAATTTGACGATCAGGTGGAAGGACTCCCAATATC 914

Db 2102 CTTTAAATATAAAATGGTCAAAATAGAAATTTGACGATCAGGTGGAAGGACTCCCAATATC 2161

QY 915 TAGCTTCTCGATATGATTTCATTTAGCTTATGATCGTGTGGGCATCCACGCGTGGTCTTATG 974

Db 2162 TAGCTTCTCGATATGATTTCATTTGACCTTAGATCGTGTGGGCATCCACGCGTGGTCTTATG 2221

QY 975 GAGGATACCTCTCCCTGATGCAATTAATGAGAGGTCAGATATCTTCAGGGTTCGTATTG 1034

Db 2222 GAGGATACCTCTCCCTGATGCAATTAATGAGAGGTCAGATATCTTCAGGGTTCGTATTG 2281

QY 1035 CTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGA 1083

Db 2282 CTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGA 2330

RESULT 3

BM557438 1042 bp mRNA linear EST 20-FEB-2002
LOCUS AGENCOURT_6578992 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:5466759
DEFINITION 5', mRNA sequence.

ACCESSION BM557438

VERSION BM557438.1 GI:18799430

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.

REFERENCE

1 (bases 1 to 1042)

NIH-MGC <http://mgc.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: DCTD/DTF

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LCM1967 row: k column: 16

High quality sequence stop: 697.

Location/Qualifiers

FEATURES

source

1..1042

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:5466759"

/tissue_type="amelanotic melanoma, cell line"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NIH_MGC_41"

/note="Organ: skin; Vector: pOTB7; Site:1: XhoI; Site:2:

EcoRI; cDNA made by oligo-dT priming. Directionally cloned

into EcoRI/XhoI sites using the following 5' adaptor:

GGCAGGAG(G). Library constructed by Ling Hong in the

Laboratory of Gerald M. Rubin (University of California,

Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and

Superscript II RT (Life Technologies). Note: this is a

NIH_MGC Library."

ORIGIN

Query Match

Best Local Similarity 72.9%; Score 789.2; DB 3; Length 1042;

Matches 794; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 280 TGAAGGAAGGATCATAGTTGATGAAGTCGAAGGCTGGTATATTTTGAAGGCACCAAGA 339

Db 18 TGGATCTAATATCAAGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGGCACCAAGA 77

QY 340 CTCCCTTTAGAGCATCACCTGTACGTPAGTCAGTACGTAATATCCCTGGAGAGGTGACAAG 399

Db 78 CTCCCTTTAGAGCATCACCTGTACGTPAGTCAGTACGTAATATCCCTGGAGAGGTGACAAG 137

QY 400 GCTGATCAGCGTGGCTACTCACATTTCTGTCGCATCAGTCAGCACTGTGACTCTTTAT 459

Db 138 GCTGATCAGCGTGGCTACTCACATTTCTGTCGCATCAGTCAGCACTGTGACTCTTTAT 197

QY 460 AAGTAAGTATAGTACCAAGAAATCCACACTGTGTCTCCCTTACAGCTATCAAGTCC 519

Db 198 AAGTAAGTATAGTACCAAGAAATCCACACTGTGTCTCCCTTACAGCTATCAAGTCC 257

QY 520 TGAAGATGACCCAACTTTGCAAAACAAAGGAATTTTGGGCCACCATTTTGGATTTCAGCAGG 579

Db 258 TGAAGATGACCCAACTTTGCAAAACAAAGGAATTTTGGGCCACCATTTTGGATTTCAGCAGG 317

QY 580 TCCTCTCTGACTATACTCTCCAGAAATTTTCTCTTTGAAAGTACTACTTGGATTTTAC 639

Db 318 TCCTCTCTGACTATACTCTCCAGAAATTTTCTCTTTGAAAGTACTACTTGGATTTTAC 377

QY 640 ATTGTATGGGATGCTCTACAAGCCTCATGATCTACAGCCTGGAAAGAAATATCTTACTGT 699

Db 378 ATTGTATGGGATGCTCTACAAGCCTCATGATCTACAGCCTGGAAAGAAATATCTTACTGT 437

QY 700 GCTGTTTATATATGCTGGTCTCTCAGGTGCTGTTGGTGAATATATCGGTTTAAAGGAGTCAA 759

Db 438 GCTGTTTATATATGCTGGTCTCTCAGGTGCTGTTGGTGAATATATCGGTTTAAAGGAGTCAA 497

QY 760 GTATTTCCGCTTGAATACCTTAGCCTCTCTAGTGTATGTTGGTGTAGTATAGACAGCAG 819

Db 498 GTATTTCCGCTTGAATACCTTAGCCTCTCTAGTGTATGTTGGTGTAGTATAGACAGCAG 557

QY 820 GGGATCCTGTCCACGAGGCTTAAATTTGAAGGCGCTTTAAATATAAAATGGGTCAAAT 879

Db 558 GGGATCCTGTCCACGAGGCTTAAATTTGAAGGCGCTTTANATATAAAATGGGTCAAAT 617

QY 880 AGAAATTCAGCATCAGGTGGAAGGACTCCAAATATCTAGCTTCTCGATATGATTCATTGA 939

Db 618 AGAAATTCAGCATCAGGTGGAAGGACTCCAAATATCTAGCTTCTCGATATGATTCATTGA 677

QY 940 CTTAGATCGTGTGGGCTATCCACGCGTGGTCTTATGAGGATACCTCTCCCTGATGGCATT 999

Db 678 CTTAGATCGTGTGGGCTATCCACGCGTGGTCTTATGAGGATACCTCTCCCTGATGGCATT 737

QY 1000 AATGCAGAGGTGAGATATCTTTCAGGGTTGCTATGTCGGGGCCCAAGTCACCTCTGTGGAT 1059

Db 738 AATGCAGAGGTGAGATATCTTTCAGGGTTGCTATGTCGGGGCCCAAGTCACCTCTGTGGAT 797

QY 1060 CTTCTATGATACAGGATACACGG 1082

Db 798 CTTCTATGATACAGGATACACGG 820

RESULT 4

BM575260

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Homidae; Homo.

REFERENCE

1 (bases 1 to 910)

AUTHORS

NIH-MGC <http://mgc.nci.nih.gov/>.

BQ675260 910 bp mRNA linear EST 15-JUL-2002
AGENCOURT_8354972 NIH_MGC_102 Homo sapiens cDNA clone IMAGE:6275478
5', mRNA sequence.

ACCESSION BQ675260

VERSION BQ675260.1 GI:21786094

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Homidae; Homo.

REFERENCE

1 (bases 1 to 910)

AUTHORS

NIH-MGC <http://mgc.nci.nih.gov/>.

Db	119	TCGATCTAATATCAAGTTGATGAAGTCAGAAGCTGGTATATTTTGAAGGCACCAAGA	178
Qy	340	CTCCCTTTAGAGCATACCTGTAGTACGTAAATCTCTGGAGAGTGACAAAG	399
Db	179	CTCCCTTTAGAGCATACCTGTAGTACGTAAATCTCTGGAGAGTGACAAAG	238
Qy	400	GCTGACTGACCGTGGCTACTCACAATTTCTGTCATCAGTCAGCAGCTGTGACTTCTTTAT	459
Db	239	GCTGACTGACCGTGGCTACTCACAATTTCTGTCATCAGTCAGCAGCTGTGACTTCTTTAT	298
Qy	460	AGATAAGTATAGTAACCAAGAGATCCACATCTGTGTCTCTTACAGCTATCAAGTCC	519
Db	299	AGATAAGTATAGTAACCAAGAGATCCACATCTGTGTCTCTTACAGCTATCAAGTCC	358
Qy	520	TGAAGATGCCCACTTCGCAAAACAAAGAAATTTTGGCCACCAATTTTGGATTACAGG	579
Db	359	TGAAGATGCCCACTTCGCAAAACAAAGAAATTTTGGCCACCAATTTTGGATTACAGG	418
Qy	580	TCCTCTTCTGACTACTCTCCAGAAATTTTCTTTTGAAGTACTACTGGATTTAC	639
Db	419	TCCTCTTCTGACTACTCTCCAGAAATTTTCTTTTGAAGTACTACTGGATTTAC	478
Qy	640	ATTGTATGGGATGCTCTCAAGCCTCATGATCTACAGCTGGAAAGAAATATCTACTGT	699
Db	479	ATTGTATGGGATGCTCTCAAGCCTCATGATCTACAGCTGGAAAGAAATATCTACTGT	538
Qy	700	GCTGTTTCATATATGCTGCTCAGGTGAGTGTGTAATCGGTTTAAAGAGTCAA	759
Db	539	GCTGTTTCATATATGCTGCTCAGGTGAGTGTGTAATCGGTTTAAAGAGTCAA	598
Qy	760	GTATTTCCGCTTGAATA-CCTAGCTCTCTAGGTATGTTGTTAGTATGATGAC-AAC	817
Db	599	GTATTTCCGCTTGAATA-CCTAGCTCTCTAGGTATGTTGTTAGTATGATGACAAAC	658
Qy	818	AGGGATCTCTGACCGAGGCTTAAATTTTGAAGCGCC--TTTAAATATAAAATGGGTC	875
Db	659	AGGGATCTCTGACCGAGGCTTAAATTTTGAAGCGCCCTTTTAAATATAAAATGGGTC	718
Qy	876	AAATAGAAATAGAGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATGATTTCA	935
Db	719	AAATAGAAATAGAGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATGATTTCA	778
Qy	936	TTGACTTTAGATCGTGTGGGATCCACGGCTGTCTCTATGAGAGTACCTCTCCCTGATG	995
Db	779	TTGACTTTAGATCGTGTGGGATCCACGGCTGTCTCTATGAGAGTACCTCTCCCTGAGG	838
Qy	996	CATTATGAGAGGTGATATCTTCAGGGTTGCTATTGCTGGGGCCCGCTCACTCTG-	1054
Db	839	CATTATGAGAGGTGATATCTTCAGGGTTGCTATTGCTGGGGCCCGCTCACTCTCGT	898
Qy	1055	TGGATCTTCTATGATACAGGATACACGGA	1083
Db	899	TGGATCTTCTATGATACAGGATACCCGGA	927
RESULT 6			
AF176779			
LOCUS			
DEFINITION			
ACCESSION			
AF176779			
AF176779.1			
VERSION			
HTC.			
KEYWORDS			
Homo sapiens (human)			
ORGANISM			
Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;			
Hominidae; Homo.			
1 (bases 1 to 1265)			
REFERENCE			
Hui, R.T., Liu, Y.Q., Wang, X.Y., Qin, B.M. and Sheng, H.			
TITLE			
Homo sapiens normal aorta mRNA MST141			
JOURNAL			
Unpublished			
2 (bases 1 to 1265)			
AUTHORS			
Hui, R.T., Liu, Y.Q., Wang, X.Y., Qin, B.M. and Sheng, H.			
TITLE			
Direct Submission			
JOURNAL			
Submitted (10-AUG-1999) Molecular Medicine Center for			
Cardiovascular Disease, Cardiovascular Institute, CAMS & PUMC, 167,			
Bei Li Shi Lu, Beijing 100037, P.R. China			
FEATURES			
Location/Qualifiers			
source			
1..1265			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/tissue_type="aorta"			
268..741			
/codon_start=1			
/product="MSTP141"			
/protein_id="AAQ13657.1"			
/db_xref="GI:33338070"			
/translation="MLYKPHDLQFEGAFYKMGQIEIDQVEGLQYLASRYDFIDLRVG			
ASLGVVYVINDNRGSGHRGLKFEQAFYKMGQIEIDQVEGLQYLASRYDFIDLRVG			
IHGWSYGVGLSLMALMQRSDFRVAIAGAPVTLWIFDYDTGTRYMGHPDQNG"			
ORIGIN			
Query Match			
Best Local Similarity			
Matches			
700; Conservative			
0; Mismatches			
0; Indels			
1; Gaps			
1; Length			
1265;			
Qy	384	CTGAGAGGTGACAAGG-CTGACTGACCGTGGCTACTCACAATTTCTGCTGCATCAGTCAG	442
Db	1	CTGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACAATTTCTGCTGCATCAGTCAG	60
Qy	443	CACGTGACCTCTTTTATAAGTATAGTAAACCAAGAGAAATCCACACTGTGTCCCTT	502
Db	61	CACGTGACCTCTTTTATAAGTATAGTAAACCAAGAGAAATCCACACTGTGTCCCTT	120
Qy	503	TACAAGCTATCAAGTCTCTCAAGTACCACTTCTTCAAGTACCACTTCTTCAAGTACCA	562
Db	121	TACAAGCTATCAAGTCTCTCAAGTACCACTTCTTCAAGTACCACTTCTTCAAGTACCA	180
Qy	563	ATTTTGGATTACAGAGTCTCTTCTTCAAGTACCACTTCTTCAAGTACCACTTCTTCAAG	622
Db	181	ATTTTGGATTACAGAGTCTCTTCTTCAAGTACCACTTCTTCAAGTACCACTTCTTCAAG	240
Qy	623	AGTACTACTCGATTTACATTTGATGAGGATGCTCTCAAGCTCATGATCTACAGCTTGA	682
Db	241	AGTACTACTCGATTTACATTTGATGAGGATGCTCTCAAGCTCATGATCTACAGCTTGA	300
Qy	683	AAGAAATATCTTACTGTCTGTTTATATATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT	742
Db	301	AAGAAATATCTTACTGTCTGTTTATATATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT	360
Qy	743	CGGTTTAAAGAGGTCAAGTATTTCCGCTTGAATATACCTAGCTCTCTAGGTTATGTTGTT	802
Db	361	CGGTTTAAAGAGGTCAAGTATTTCCGCTTGAATATACCTAGCTCTCTAGGTTATGTTGTT	420
Qy	803	GTAGTGATAGACAAACAGGGGATCTCTTCAAGCTGAGGCTTAAATTTGAAGGCGCTTTAA	862
Db	421	GTAGTGATAGACAAACAGGGGATCTCTTCAAGCTGAGGCTTAAATTTGAAGGCGCTTTAA	480
Qy	863	TATAAATGGGTCAAAATAGAAATTTGACATCAGGTGGAAGGACTCCAATATCTAGCTTCT	922
Db	481	TATAAATGGGTCAAAATAGAAATTTGACATCAGGTGGAAGGACTCCAATATCTAGCTTCT	540
Qy	923	CGATATGATTTTCAATTTGATCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT	982
Db	541	CGATATGATTTTCAATTTGATCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT	600
Qy	983	CTCTCCCTGATGGCATTAATGACAGGTCTCAGATATCTTCAAGGTTGCTTATTTCTGGGCC	1042
Db	601	CTCTCCCTGATGGCATTAATGACAGGTCTCAGATATCTTCAAGGTTGCTTATTTCTGGGCC	660
Qy	1043	CCAGTCACTCTGTGGATCTTCTTATGATACAGGATACACCGA	1083
Db	661	CCAGTCACTCTGTGGATCTTCTTATGATACAGGATACACCGA	701

```
RESULT 7
AY411617
LOCUS
DEFINITION
AY411617
ACCESSION
AY411617.1
KEYWORDS
GSS
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
REFERENCE
AUTHORS
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL
Science 302 (5652), 1960-1963 (2003)
PUBMED
14671302
REFERENCE
AUTHORS
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE
Direct Submission
JOURNAL
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT
This sequence was made by sequencing genomic exons and ordering
them based on alignment
FEATURES
Source
Location/Qualifiers
1..2634
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1..>2634
/gene="PPP8"
/locus_tag="HCM4283"
ORIGIN
Query Match 61.9%; Score 670.6; DB 10; Length 2634;
Best Local Similarity 81.9%; Pred. No. 2.4e-179;
Matches 718; Conservative 0; Mismatches 159; Indels 0; Gaps 0;
Qy 207 CATTCCGTTATCTTAACAGAGTACAGCAATCTCTAAAGTCATTTTAAGATGTCAGAAA 266
Db 1457 CAANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1516
Qy 267 TAATGATTGATGCTGAAGAGGATCATAGTTGATGAAGTCAGAAAGCTGGTATATTG 326
Db 1517 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1576
Qy 327 AAGGCCAACAGACTCCCTTTAGAGCATCACCTGTAGTACGTTCAGTTAGTAAATCCCTG 386
Db 1577 AAGGCCAACAGACTCTCTTTGGAGCATCACCTGTAGTACGTTCAGTTAGTAAATCCCTG 1636
Qy 387 GAGAGGTGACAAAGGCTGACTACCGTGGCTACTCACATTCTCTGTCATCAGTCAGCACT 446
Db 1637 GAGAAGTGGTGAAGGCTGACTACCGTGGCTACTCACACTCTCTGCTCCCTCAGCCGCAT 1696
Qy 447 GTGACTTCTTTAAGTAAGTATAGTACAGAGATCCACACTGCTGTGCTCCCTTACA 506
Db 1697 GTGACTTCTTTAAGTAAGTATAGTACAGAGATCCACACTGCTGTGCTCCCTTACA 1756
Qy 507 AGCTATCAAGTCTCAAGTACCACTTCGCAACAGAGGAAATTTGGGACAGATT 566
Db 1757 AACTCTCAAGTCTGAGGATCAAGTATCAAAACAGAGGAAATTTGGGACAGATT 1816
Qy 567 TGGATTACAGAGTCTCTCTGACTACTCTCTCAGAAATTTCTCTTTTGAAGTA 626
Db 11076861
```

```
1817 TGGATTACAGAGTCTCTCTGACTACACCCCTCCAGAAATTTTCTTTTGAAGTA 1876
Qy 627 CTACTGGATTTACATTCATGGGATGCTTACAGGCTCATGATCTACAGCCCTGGAAGA 686
Db 1877 CTACTGGATTTACATTCATGGGATGCTTACAGGCTCATGATCTACAGCCCTGGAAGA 1936
Qy 687 AATATCTCTACTGTGCTCTTTCATATATGGTGGTCTCAGGTGAGTGGTGAATAATCGT 746
Db 1937 AATATCTCTACTGTGCTCTTTCATATATGGTGGTCTCAGGTGAGTGGTGAATAATCGT 1996
Qy 747 TTAAGAGAGTCAAGTATTTCCGCTTGAATACCTAGCTCTCTAGTGTATGTGTTGTAG 806
Db 1997 TTAAGAGAGTCAAGTATTTCCGCTTGAATACCTAGCTCTCTAGTGTATGTGTTGTAG 2056
Qy 807 TGATAGACAAACAGGGGATCTGTACCGAGGGCTTAAATTTCAAGGCGCCCTTAAATATA 866
Db 2057 TGATAGACAAACAGGGGATCTGTACCGAGGGCTTAAATTTCAAGGCGCCCTTAAATATA 2116
Qy 867 AAATGGGTCAAAATAGAAATTTGACGATCAGGTGGAGGACTCCAATATCTAGCTTCGAT 926
Db 2117 AAATGGGTCAAAATAGAAATTTGACGATCAGGTGGAGGACTCCAATATCTAGCTTCGAT 2176
Qy 927 ATGATTTTCATTGACTTAGATCGTGGGATCCAGGCTGGTCTCTATCGAGGATACCTCT 986
Db 2177 ATGATTTTCATTGACTTAGATCGTGGGATCCAGGCTGGTCTCTATCGAGGATACCTCT 2236
Qy 987 CCCTGATGGCATTAATGCGAGAGTTCAGATATCTTCAGGGTGGCTATTGCTGGGCCCCAG 1046
Db 2237 CCCTGATGGCATTAATGCGAGAGTTCAGATATCTTCAGGGTGGCTATTGCTGGGCCCCAG 2296
Qy 1047 TCATCTGTGGATCTTCTATGATACAGGATACACGGA 1083
Db 2297 TCACCTGTGGATCTTCTATGATACAGGATACACGGA 2333
RESULT 8
AK016546
LOCUS
DEFINITION
AK016546
ACCESSION
AK016546.1
VERSION
AK016546.1
KEYWORDS
HTC; CAP trapper.
SOURCE
Mus musculus
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1
AUTHORS
Carninci, P. and Hayashizaki, Y.
TITLE
High-efficiency full-length cDNA cloning
JOURNAL
Meth. Enzymol. 303, 19-44 (1999)
PUBMED
10349636
REFERENCE
2
AUTHORS
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL
Genome Res. 10 (10), 1617-1630 (2000)
PUBMED
11042159
REFERENCE
3
AUTHORS
Shibata, K., Itoh, M., Aizawa, K., Nagasaka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Okada, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Yoneda, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
TITLE
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
JOURNAL
Genome Res. 10 (11), 1757-1771 (2000)
PUBMED
11076861
```


RESULT 9
AK029788
LOCUS

DEFINITION

AK029788 5517 bp mRNA linear HTC 03-APR-2004
Mus musculus adult male testis cDNA, RIKEN full-length enriched
library, clone:4930560C15 product:DIPEPTIDYL PEPTIDASE 8 homolog
[Homo sapiens], full insert sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AK029788 5517 bp mRNA linear HTC 03-APR-2004
Mus musculus adult male testis cDNA, RIKEN full-length enriched
library, clone:4930560C15 product:DIPEPTIDYL PEPTIDASE 8 homolog
[Homo sapiens], full insert sequence.

AK029788 1 GI:26081520
HTC; CAP trapper.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.

1 Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
10349636

2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
11042159

3 Shibata, K., Itoh, M., Aizawa, K., Nagao, K., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujimaki, S., Inoue, K., Ozawa, Y., Izawa, M., Ohara, S., Watanabe, M.,
Itoneda, Y., Ishikawa, T., Tanaka, K., Tanaka, T., Matsuda, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
11076861

4 The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)

5 The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)

6 (bases 1 to 5517)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kejima, Y., Kondo, S., Kono, H., Kouda, M.,
Koye, S., Kurihara, C., Maenuyama, T., Miyazaki, A., Murata, Y.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohnato, N.,
Okazaki, Y., Saito, K., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.
Direct Submission

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gs.riken.jp,
URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)

7 Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gs.riken.jp,
URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)

8 Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gs.riken.jp,
URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)

9 Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gs.riken.jp,
URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)

Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/.

FEATURES

source

1..5517
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM DB:4930560C15"
/db_xref="taxon:10090"
/clone="4930560C15"
/sex="male"
/tissue_type="testis"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
221..4837
/note="DIPEPTIDYL PEPTIDASE 8 homolog [Homo sapiens]
(SPTR|Q9HBM5, evidence FASTV, 95.7%ID, 100%length,
match=2649)
putative"

misc_feature

ORIGIN

Query Match 61.8%; Score 669.6; DB 4; Length 5517;
Best Local Similarity 89.6%; Pred. No. 5.9e-179;
Matches 720; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

QY 280 TCAAGGAGGATCATCTGTCAGTCAGAGGCTGGTATATTTTGAAGGCCACCAAGA 339
DB 3715 TGGATCTAATATCTGGTTCATGAGCCAGAAAGCTGGTGACTTTGAAGGCCACCAAGA 3774
QY 340 CTCCTTTAGAGCATCCTCTGACGTAGTCAGTACGTAAATCTCTGGAGAGTGACAG 399
DB 3775 CTCCTTTGAGCATCCTCTGACGTAGTCAGTACGTAAATCTCTGGAGAGTGAG 3834
QY 400 GCTGATGACCGTGGCTACTCACATTTCTGCTGCATGACGTAGTCAGTACGTGACTTCTTAT 459
DB 3835 GCTGATGACCGTGGCTACTCACATTTCTGCTGCATGACGTAGTCAGTACGTGACTTCTTAT 3894
QY 460 AAGTAAGTATAGTAACCAAGAGAAATTTGGGCCACCATTTTGGGATTCAGCAGG 519
DB 3895 AAGTAAGTATAGTAACCAAGAGAAATTTGGGCCACCATTTTGGGATTCAGCAGG 3954
QY 520 TGAAGATGACCAACTTGCAGAAATTTTGGGCCACCATTTTGGGATTCAGCAGG 579
DB 3955 TGAAGATGACCAACTTGCAGAAATTTTGGGCCACCATTTTGGGATTCAGCAGG 4014
QY 580 TCCTCTCTGACTATATCTCTCAGAAATTTTCTCTTTGAAAGTACTACTGGATTAC 639
DB 4015 TCCTCTCTGACTATATCTCTCAGAAATTTTCTCTTTGAAAGTACTACTGGATTAC 4074
QY 640 ATTGTATGGGATGCTCTACAGGCTCATGATCTACAGCTGTGCTCTTACAGCTATCAAGTCC 699
DB 4075 ACTGTATGGGATGCTCTACAGGCTCATGATCTACAGCTGTGCTCTTACAGCTATCAAGTCC 4134
QY 700 GCTGTTCATATATGCTGCTCTCAGTGGAGTGGTGAATATCGTTTAAAGGATCNA 759
DB 4135 GTTATTATATATGCTGCTCTCAGTGGAGTGGTGAATATCGTTTAAAGGATCNA 4194
QY 760 GTTATTTCGCTTGAATACCTTACGCTCTCTAGGTTATGCTGTTAGTATGATGACACAG 819
DB 4195 GTTATTTCGCTTGAATACCTTACGCTCTCTAGGTTATGCTGTTAGTATGATGACACAG 4254
QY 820 GGGATCTGTCTCAGCGGCTTAAATTTGAAGGCGCTTTAAATATAAATGGTCAAT 879
DB 4255 GGGATCTGTCTCAGCGGCTTAAATTTGAAGGCGCTTTAAATATAAATGGTCAAT 4314
QY 880 AGAAATGACGATCAGGTGAAGGACTCCCAATATCTAGCTTCTCGATGATGATTTCATGGA 939
DB 4315 AGAAATGACGATCAGGTGAAGGACTCCCAATATCTAGCTTCTCGATGATGATTTCATGGA 4374

```
QY 940 CTTAGATCGTGGGCATCCAGCGTGGTCCCTATGAGGATACCTCTCCCTGATGCATT 999
Db 4375 CTTGGATCAGTGGGCATCCAGCGTGGTCCCTATGAGGATACCTCTCCCTGATGCATT 4434
QY 1000 AATGACAGGCTCAGATATCTTTCCAGGGTTGCTATTGCTGGGGCCCAAGTCACTCTGTTGGAT 1059
Db 4435 AATGACAGGCTCGATATCTTTCCGGGTGGCTATTGCTGGGGCCCAAGTCACTCTGTTGGAT 4494
QY 1060 CTTCTATGATACAGGATACAGGA 1083
Db 4495 CTTCTATGATACAGGATACAGGA 4518

RESULT 10
BQ671635
LOCUS BQ671635
DEFINITION AGENCOURT_8203887 NIH_MGC_102 Homo sapiens cDNA clone IMAGE:6255645
5', mRNA sequence.
ACCESSION BQ671635
VERSION BQ671635.1 GI:21782469
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
REFERENCE 1 (bases 1 to 968)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Cloning by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCW2407 row: a column: 22
High quality sequence stop: 630.

FEATURES
source
1..968
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6255645"
/tissue_type="epidermoid carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_102"
/note="Organ: salivary gland; Vector: pOT87; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed
by Ling Hong in the laboratory of Gerald M. Rubin
(University of California, Berkeley) using ZAP-cDNA
synthesis kit (Stratagene) and Superscript II RT (Life
Technologies). Note: this is a NIH_MGC Library."

ORIGIN
Query Match 61.4%; Score 665.2; DB 5; Length 968;
Best Local Similarity 97.6%; Pred. No. 5.Be-178;
Matches 696; Conservative 0; Mismatches 14; Indels 3; Gaps 2;
QY 280 TGAAGGAGGATCATAGTTGATGAAGTCAGAGCGCTGATATTTTGAAGGCCACCAAGA 339
Db 107 TGGATCTATATCCCAAGTTGATGAAGTCAGAGCGCTGATATTTTGAAGGCCACCAAGA 166
QY 340 CTCCTCTTTAGAGCATCACCTGTACGTAGTCAGTTAGTAAATCTCTGGAGAGGTGACAAG 399
Db 167 CTCCTCTTTAGAGCATCACCTGTACGTAGTCAGTTAGTAAATCTCTGGAGAGGTGACAAG 226
QY 400 GCTGACTGACCGTGGCTACTCACATTTCTTGTGTCATCAGTCAGCACTGTGACTTTCTTAT 459
```

```
Db 227 GCTGACTGACCGTGGCTACTCACATTTCTTGTGTCATCAGTCAGCACTGTGACTTTCTTAT 286
QY 460 AAGTAAGTATAGTAACACAGAGAAATCCACACTGTGTGCTCTTTTACAAGCTATCAAGTCC 519
Db 287 AAGTAAGTATAGTAACACAGAGAAATCCACACTGTGTGCTCTTTTACAAGCTATCAAGTCC 346
QY 520 TGAAGATGACCAACTTTCGAAAAAACAAGGAAATTTTGGGGCCCAACATTTTGGATTTCAGCAGG 579
Db 347 TGAAGATGACCAACTTTCGAAAAAACAAGGAAATTTTGGGGCCCAACATTTTGGATTTCAGCAGG 406
QY 580 TCCTCTTCCTGACTATACCTCCAGAGAAATTTCTCTTTTGAAGTACTACTCGATTTC 639
Db 407 TCCTCTTCCTGACTATACCTCCAGAGAAATTTCTCTTTTGAAGTACTACTCGATTTC 466
QY 640 ATTGATGCGATGCTCTTACAAGCCTCATGATCTACAGCCTGGAAAGAAATATCTACTGT 699
Db 467 ATTGATGCGATGCTCTTACAAGCCTCATGATCTACAGCCTGGAAAGAAATATCTACTGT 526
QY 700 GCTGTTTCATATATGTTGCTCTCAGGTGTCAGTGGTGAATAATCGGTTTAAAGGAGTCAA 759
Db 527 GCTGTTTCATATATGTTGCTCTCAGGTGTCAGTGGTGAATAATCGGTTTAAAGGAGTCAA 586
QY 760 GTATTTCGCTTGAATACCTAGCTCTCTAGGTTATGCTGTTCTAGTGATGATACACAG 819
Db 587 GTATTTCGCTTGAATACCTAGCTCTCTAGGTTATGCTGTTCTAGTGATGATACACAG 646
QY 820 GGGATCCTGTCTACCGAGGGCTTAAATTTGAAGGGCGCTTTAAATATATAAAATGGTCAAAAT 879
Db 647 GGGATCCTGTCTACCGAGGGCTTAAATTTGAAGGGCGCTTTANATATATAAAATGGTCAAAAT 706
QY 880 AGAAATGACGATCAGGTGGAAGGACTCCAAATATCTAGCTTCTCGATATGATTTCATTGA 939
Db 707 AGAAATGACGATCAGGTGGAAGGACTCCAAATATCTAGCTTCTCGATATGATTTCATTGA 766
QY 940 CTTAGATC-GTCTGGGCATCCACCGCTGGTCC--TATGGAGGATACCTCTCCC 989
Db 767 CTTAGATC-GTCTGGGCATCCACCGCTGGCCCTATGGGAGGATACCTCTCCC 819

RESULT 11
LOCUS BQ671635
DEFINITION 602869453F1 NIH_MGC_102 Homo sapiens cDNA clone IMAGE:5013996 5',
mRNA sequence.
ACCESSION BQ671635
VERSION BQ671635.1 GI:14502420
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
REFERENCE 1 (bases 1 to 1041)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: CLONETECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCW1820 row: j column: 13
High quality sequence stop: 847.

FEATURES
source
1..1041
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"

ORIGIN
```

```
/clones="IMAGE:5013996"
/tissue_type="epidermoid carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_102"
/notes="Organ: salivary gland; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed
by Ling Hong in the laboratory of Gerald M. Rubin
(University of California, Berkeley) using ZAP-cDNA
synthesis kit (Stratagene) and Superscript II RT (Life
Technologies). Note: this is a NIH_MGC Library."

ORIGIN
Query Match 61.1%; Score 662.2; DB 2; Length 1041;
Best Local Similarity 96.4%; Pred. No. 4.3e-177;
Matches 784; Conservative 0; Mismatches 18; Indels 11; Gaps 10;

QY 280 TGAAGGAGGATCATAGTCAAGTCAAGGCTGGTATATTTTGAAGCACCACGA 339
Db 110 TGGATCTATATCCAGTGTATGAAGTCAGAGGCTGGTATATTTTGAAGCACCACGA 169

QY 340 CTCCTCTTTAGAGCATCACTGTACGTAGTACGTTACGTAATCTCTGGAGAGGTGACAA 399
Db 170 CTCCTCTTTAGAGCATCACTGTACGTAGTACGTTACGTAATCTCTGGAGAGGTGACAA 229

QY 400 GCTGACTGACCGTGTACTCACATTTCTGTCGATCAGTCAGCACTGTGACTTCTTTAT 459
Db 230 GCTGACTGACCGTGTACTCACATTTCTGTCGATCAGTCAGCACTGTGACTTCTTTAT 289

QY 460 AAGTAAGTATAGTAAACAGAGAAATCCACACTGTGTCTCTTTTCAAGCTATCAAGTCC 519
Db 290 AAGTAAGTATAGTAAACAGAGAAATCCACACTGTGTCTCTTTTCAAGCTATCAAGTCC 349

QY 520 TGAAGATGACCAACTTGTCAAAACAAAGGAATTTTGGGCCACCATTTGGATTTCAGCAG 579
Db 350 TGAAGATGACCAACTTGTCAAAACAAAGGAATTTTGGGCCACCATTTGGATTTCAGCAG 409

QY 580 TCCTCTCTCAGCTATCTCTCCAGAAATTTCTCTTTTGAAGTACTACTGGAATTTAC 639
Db 410 TCCTCTCTCAGCTATCTCTCCAGAAATTTCTCTTTTGAAGTACTACTGGAATTTAC 469

QY 640 ATTGTATGGGATGCTCTACAGCCTCATGATCTACAGCCTGGAAAGAAATATCTTACTGT 699
Db 470 ATTGTATGGGATGCTCTACAGCCTCATGATCTACAGCCTGGAAAGAAATATCTTACTGT 529

QY 700 GCTGTTTATATATGCTGCTCTCAGGTGCAATTTGGTGAATATCGTTTA--AAGGATC 757
Db 530 GCTGTTTATATATGCTGCTCTCAGGTGCAATTTGGTGAATATCGTTTACAGCGATC 589

QY 758 AAGTATTTCCGCTTGAATACCCCTAGCCTCTTAGGTTATGCTGTTAGTATAGACAA 817
Db 590 AAGTATTTCCGCTTGAATACCCCTAGCCTCTTAGGTTATGCTGTTAGTATAGACAA 649

QY 818 AGGGATCCTGTACCGAGGCTTAAATTTGAAGGCGCTTTAAATATAAATAGGTC- 877
Db 650 AGGGATCCTGTACCGAGGCTTAAATTTGAAGGCGCTTTTAAATATAAATAGGTC-A 707

QY 878 ATAGAAATATGACATCAGGTGGAAGGACTCCAA--TATCTAGCTTCTCGATATGATTTCA- 935
Db 708 ATAGAAATATGACATCAGGTGGAAGGACTCCAAATTTCTAGCTTCTCGATATGATTTCA 767

QY 936 TTGACTTAGATCGTGTGGGATCCACCGCTGGTCTTATGGA--GGATACCTCTCCCTG-AT 993
Db 768 TTGACTTAGATCGTGTGGGATCCACCGCTGGTCTTATGGA--GGATACCTCTCCCTGAA 827

QY 994 GGCATTAATGACAGGTCAGATA--TCCTCAGGGTGTCTATTGCT--GGGCCCCCAGTCACT 1051
Db 828 GGCATTAATGACAGGTCAGATACTCTTCAGGGTGCATATAGCTGGGGCCCCCAGTCACT 887

QY 1052 CTGTGGATCTTC--TATGATACAGGATACACGGA 1083
Db 888 CTGTGGATCTTCCTTATGATACAGGATACACGGA 920
```

RESULT 12
CN427231

LOCUS 17000455488465 GRN_EB Homo sapiens cDNA 5', mRNA linear EST 16-MAY-2004
DEFINITION CN427231
ACCESSION CN427231
VERSION CN427231.1 GI:47414825
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE
AUTHORS

1 (bases 1 to 726)
Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G.J.,
Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M.S., Mandalam, R.,
Lebkowski, J. and Stanton, L.W.
Transcriptome characterization elucidates signaling networks that
control human ES cell growth and differentiation
Nat Biotechnol. 22 (6), 707-716 (2004)

JOURNAL
PUBMED

15146197

COMMENT

Contact: Brandenberger R
Regenerative Medicine
Geron Corporation
230 Constitution Drive, Menlo Park, CA 94025, USA
Tel: 650 473 8658
Fax: 650 473 7760
Email: rbrandenberger@geron.com
Insert Length: 726 Std Error: 0.00.

FEATURES
source

1..726
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="embryonic stem cells, embryoid bodies
derived from H1, H7 and H9 cells"
/clone_lib="GRN_EB"
/note="oligo dT primed, full-length enriched cDNA library
from embryoid body outgrowths derived from hES cell lines
H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free
conditions."

ORIGIN

Query Match 60.9%; Score 659.8; DB 7; Length 726;
Best Local Similarity 99.4%; Pred. No. 1.8e-176;
Matches 683; Conservative 0; Mismatches 2; Indels 2; Gaps 2;
QY 399 GGCTGACTGACCGTGGCTACTCATTTCTTGTGCATCAGTCAGCACTGTGACTTCTTTA 458
Db 2 GGCTGACTGACCGTGGCTACTCATTTCTTGTGCATCAGTCAGCACTGTGACTTCTTTA 61
QY 459 TAAGTAAGTATAGTAACAGAGAAATCCCACTGTGTCTTTCCTTACAGCTATCAAGTC 518
Db 62 TAAGTAAGTATAGTAACAGAGAAATCCCACTGTGTCTTTCCTTACAGCTATCAAGTC 121
QY 519 CTGAAGATGACCCAACTTGGCAAAACAAAGAAATTTGGGCCACCATTTGGATTTCAGCG 578
Db 122 CTGAAGATGACCCAACTTGGCAAAACAAAGAAATTTGGGCCACCATTTGGATTTCAGCG 181
QY 579 GTCTCTTCTGACTATACCTCCAGAAATTTCTTTTGAAGTACTACTGGAATTTA 638
Db 182 GTCTCTTCTGACTATACCTCCAGAAATTTCTTTTGAAGTACTACTGGAATTTA 241
QY 639 CATTTATGGGATGCTCTACAAGCCTCATGATCTACAGCCTGGAAAGAAATATCTTACTG 698
Db 242 CATTTATGGGATGCTCTACAAGCCTCATGATCTACAGCCTGGAAAGAAATATCTTACTG 301
QY 699 TGCTGTTTCATATAT--GGTGGTCTCTCAGGTGCAAGTATCGTGAATATCGTTAAAGAGTC 757
Db 302 TGCTGTTTCATATATGGGTGGTCTCTCAGGTGCAAGTATCGTGAATATCGTTAAAGAGTC 361
QY 758 AAGTATTTCCGCTTGAATATCCCTAGCCTCTCTAGTTATGTTGTTAGTATAGACAA 817

||||| 362 AAGTATTTCCGCTTGAATACCCCTAGCCTCTCTAGGTATGTGGTTAGTGATAGACAAC 421
Db
||||| 818 AGGGGATCTCTGACCGAGGCTTAAATTTGAAGCGCTTTAAATATAAATAGGTCAC 877
Qy
Db 422 AGGGGATCTCTGACCGAGGCTTAAATTTGAAGCGCTTTAAATATAAATAGGTCAC 481
Qy 878 ATAGAAATTCAGCATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGGATATGATTTCAAT 937
Db 482 ATAGAAATTCAGCATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGGATATGATTTCAAT 541
Qy 938 GACTTAGATCGTGTGGGATCCACCGCTGGTCTTATGAGGATACCTCTCCCTGATGGCA 997
Db 542 GACTTAGATCGTGTGGGATCCACCGCTGGTCTTATGAGGATACCTCTCCCTGATGGCA 601
Qy 998 TTAATG-CAGAGGTCAGATATCTCAGGCTGCTATGCTGGGCCCCAGTCACCTCTGTG 1056
Db 602 TTAATG-CAGAGGTCAGATATCTCAGGCTGCTATGCTGGGCCCCAGTCACCTCTGGG 661
Qy 1057 GATCTTCTATGATACAGGATACACGGA 1083
Db 662 GATCTTCTATGATACAGGATACACGGA 688

RESULT 13
AF175225 1292 bp mRNA linear HTC 01-AUG-2003
LOCUS Homo sapiens tissue-type aorta MSRP135 mRNA, complete cds.
DEFINITION AF175225
ACCESSION AF175225.1 GI:33338055
VERSION
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 1292)
AUTHORS Zhao, B., Xu, H.S., Tong, Y.K., Sheng, H., Qin, B.M., Liu, Y.Q., Liu, B.,
Wang, X.Y., Zhang, Q., Song, L., Gao, Y., Zhang, C.L., Ye, J., Ji, X.J.,
Liu, B.H., Lu, H., Chen, J.Z., Cai, M.Q., Zheng, W.Y., Teng, C.Y.,
Liu, Q., Yu, L.T., Lin, J., Gong, Q., Zhang, A.M., Gao, R.L. and Hui, R.T.
Direct Submission
TITLE Submitted (04-AUG-1999) Molecular Medicine Center for
JOURNAL Cardiovascular Disease, Cardiovascular Institute, CAMS & PUMC, 167,
Bei Li Shi Lu, Beijing 100037, P.R. China
FEATURES
source Location/Qualifiers
1..1292
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="aorta"
142..1023
/codon_start=1
/product="MSTP135"
/protein_id="AAQ13650.1"
/db_xref="GI:33338056"
/translation="MTOLAKORNGPFWIQVLFITLLOKFSLLKVLDLHQMCGS
TSRSGSTGLERNILLXFIYGGKPKVOLNNRKGKVFRLATLPLGVVXVGINRGS
CPRLKGFEPKFMGQIIDDVQGLQIYASRFDILDRVGHWSYGGYLSMAL
MQRSDIFVAIAGAPVTLFIYDTGYRMGHPDQNGYILGVSVMQAKFPSEPN
RLLLHGFLENVHFAHTSILLSLVLRAGKPYDLQIYQERHSIRVPESGHEHLHL
HYLENLGSRIALIKVI"

ORIGIN
Query Match 59.3%; Score 642; DB 4; Length 1292;
Best Local Similarity 96.3%; Pred. No. 2.6e-171;
Matches 678; Conservative 0; Mismatches 22; Indels 4; Gaps 2;

Qy 384 CTGGAGAGTCAAGGCTGACTGACCGTGGCTACTCACATCTTCTGCTCATCAGTCAGC 443
Db 1 CTGGAGAGTCAAGGCTGACTGACCGTGGCTACTCACATCTTCTGCTCATCAGTCAGC 60
Qy 444 ACTGTGACTCTTTTATAAGTAAGTATAGTAACAGAGAAGATCCACACTGTGTGCTCCCTTT 503

||||| 61 ACTGTGACTCTTTTATAAGTAAGTATAGTAACAGAGAAGATCCACACTGTGTGCTCCCTTT 120
Db
Qy 504 ACAAGCTTATCAAGTCTCTGAAGATGACCCAACTTTGCAAAACAAAGAAATTTTGGGCCACCA 563
Db 121 ACAAGCTTATCAAGTCTCTGAAGATGACCCAACTTTGCAAAACAAAGAAATTTTGGGCCACCA 180
Qy 564 TTTTGGGATTCAGCAGGTCCTCTTCTCGATATACTCTCTCCAGAAATTTTCTCTTTTGA 623
Db 181 TTTTGGGATTCAGCAGGTCCTCTTCTCGATATACTCTCTCCAGAAATTTTCTCTTTTGA 240
Qy 624 GTACTACTGGATTTTACATCTGATGGGATGCTCTACAGGCTCATGATCTACACA--GCCTG 680
Db 241 GTACTACTGGATTTTACATCTGATGGGATGCTCTACAGGCTCATGATCTACAGGCTG 300
Qy 681 GAAAGAAATATCTTACTGT--GCTGTTTCAATATATGTTGGTCTCTCAGTGCAGTTGGTGAAT 739
Db 301 GAAAGAAATATCTTACTGNGGCTGTTTCAATATATGTTGGGCTCTTAAAGTGCAGTTGGTGAAT 360
Qy 740 AATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTTACCTAGCTCTCTAGGTTATGTG 799
Db 361 AATCGGTTTAAAGGAGTCAAGTATTTTCCGCTTGAATACCTTACCTTCTTTAGGTTATGTG 420
Qy 800 GTTGTAGTGTAGACACAGGGGATCTGTCCACCGAGGGCTTAAATTTTGAAGCGCTTT 859
Db 421 GTTGNAGGATAGACACAGGGGATCTGTCCCGAGGGCTTAAATTTTGAAGCGCTTT 480
Qy 860 AAATATAAATATGGGTCAAATAGAAATTTGACGATCAGGTGGAAGGACTCCCAATATCTAGCT 919
Db 481 AAATATAAATATGGGTCAAATAGAAATTTGACGATCAGGTGGAAGGACTCCCAATATCTAGCT 540
Qy 920 TCTCGATATGATTTTCAATGACTTAGATCTGTTGGGATCCACGGCTGTGCTTATGAGGA 979
Db 541 TCTCGATATGATTTTCAATGACTTAGATCTGTTGGGATCCACGGCTGTGCTTATGAGGA 600
Qy 980 TACCTCTCCCTGATGGCATTAAATGACAGGTCAGATATCTTACAGGCTTCTATGCTGG 1039
Db 601 TACCTCTCCCTGATGGCATTAAATGACAGGTCAGATATCTTACAGGCTTCTATGCTGG 660
Qy 1040 GCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGA 1083
Db 661 GCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGA 704

RESULT 14
BE888665 674 bp mRNA linear EST 20-OCT-2000
LOCUS 601513061F1 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:3914582 5',
DEFINITION mRNA sequence.
ACCESSION BE888665
VERSION BE888665.1 GI:10345197
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 674)
AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM9736 row: i column: 15
High quality sequence stop: 662.
Location/Qualifiers

FEATURES

```
source
1. .674
/organism="Homo sapiens"
/mol_type="mrna"
/db_xref="taxon:9606"
/clone="IMAGE:3914582"
/tissue_type="leiomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_71"
/notes="Organ: uterus; Vector: PCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 2.1 kb. "

ORIGIN
Query Match 57.2%; Score 619.8; DB 2; Length 674;
Best Local Similarity 98.5%; Pred. No. 4.5e-165;
Matches 657; Conservative 0; Mismatches 7; Indels 3; Gaps 3;
QY 343 CCCTTTAGAGCATCACCCTGACGTAGTCAGTTACGTAAATCCTGGAGAGGTGACAAGGCT 402
Db 1 CCCTTTAGAGCATCACCCTGACGTAGTCAGTTACGTAAATCCTGGAGAGGTGACAAGGCT 60
QY 403 GACTGACCGTGGCTACTCACATTTCTGCTGCATCAGTCAGCACTGTGACTTCTTTATAAG 462
Db 61 GACTGACCGTGGCTACTCACATTTCTGCTGCATCAGTCAGCACTGTGACTTCTTTATAAG 120
QY 463 TAAGTATAGTAACACAGAGAATCCACACTGTGTGCCCTTTTACAAGCTATCAAGTCTCTGA 522
Db 121 TAAGTATAGTAACACAGAGAATCCACACTGTGTGTGCCCTTTTACAAGCTATCAAGTCTCTGA 180
QY 523 AGATGACCCAACTTGCACCAAGGAATTTTGGGCCACCATTTTGGATTTCAGCAGGTCC 582
Db 181 AGATGACCCAACTTGCACCAAGGAATTTTGGGCCACCATTTTGGATTTCAGCAGGTCC 240
QY 593 TCTTCTTGACTATACCTCCAGAAATTTCTCTTTTGAAGTACTACTGATTTACATT 642
Db 241 TCTTCTTGACTATACCTCCAGAAATTTCTCTTTTGAAGTACTACTGATTTACATT 300
QY 643 GTATGGGATGCTCTACAGCTCATGATCTACAGCTGGAAGAAATATCTACTGTGCT 702
Db 301 GTATGGGATGCTCTACAGCTCATGATCTACAGCTGGAAGAAATATCTACTGTGCT 360
QY 703 GTTCATATATGGTGGTCTCAGGTGCAATTTTGAAGTAAATCGTTTAAAGGAGTCAAGTA 762
Db 361 GTTCATATATGGTGGTCTCAGGTGCAATTTTGAAGTAAATCGTTTAAAGGAGTCAAGTA 420
QY 763 TTTCCGCTTGAATACCCCTAGCCTCTCTAGGTTATGTGGTTGTAGTGATAGACACAGGGG 822
Db 421 TTTCCGCTTGAATACCCCTAGCCTCTCTAGGTTATGTGGTTGTAGTGATAGACACAGGGG 480
QY 823 ATCTGTACCCGAGGGCTTAAATTTGAAGGCGCCTTTAAATATATAAATGGGTCAACATAGA 882
Db 481 ATCTGTACCCGAGGGCTTAAATTTGAAGGCGCCTTTAAATATATAAATGGGTCAACATAGA 540
QY 883 AATTGACGATCAGTGGAGG-ACTCCAATATCTAGCTTCTCGATATGATTTTCATTGACT 941
Db 541 AATTGACGATCAGTGGAGGAACTCCAATATCTAGCTTCTCGATATGATTTTCATTGACT 600
QY 942 TAGATCTGTGGGCATCCAGGCTGTCTATGGAGGATACCTCTCCCTGATGCATTA 1001
Db 601 TAGATCTGTGGGCATCCAGG- TGGTCTTATGGAGGATA-CTCTCCCTGATGGCATTA 658
QY 1002 TGCAGAG 1008
Db 659 TGCNAAG 665

RESULT 15
BGI65205
LOCUS 602344074F1 NIH_MGC_89 Homo sapiens cdna clone IMAGE:4453987 5',
DEFINITION mRNA sequence.
ACCESSION BGI65205
VERSION BGI65205.1 GI:12671908
```

```
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1. (bases 1 to 639)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strauberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: ATCC
cdna Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10245 row: d column: 20
High quality sequence stop: 639.
FEATURES
Location/Qualifiers
1. 639
/organism="Homo sapiens"
/mol_type="mrna"
/db_xref="taxon:9606"
/clone="IMAGE:4453987"
/tissue_type="hyponephroma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_89"
/notes="Organ: Kidney; Vector: PCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally; oligo-dt primed. Average insert size 1.3 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."
ORIGIN
Query Match 56.9%; Score 616; DB 2; Length 639;
Best Local Similarity 99.7%; Pred. No. 5.4e-164;
Matches 638; Conservative 0; Mismatches 0; Indels 2; Gaps 2;
QY 418 CTCACATTCCTGTGTCATCAGTCAGCACTGTGACTTCTTTTATAAGTATAGTAAACCA 477
Db 1 CTCACATTCCTGTGTCATCAGTCAGCACTGTGACTTCTTTTATAAGTATAGTAAACCA 60
QY 478 GAAGAAATCCACACTGTGTGCCCTTTTACAAGCTATCAAGTCTGGAAGATGACCCAACTTG 537
Db 61 GAAGAAATCCACACTGTGTGCCCTTTTACAAGCTATCAAGTCTGGAAGATGACCCAACTTG 120
QY 538 CAAACAAAGGAATTTTGGGCCACCATTTTGGATTTCAGCAGGTCTCTTCTGACTATAC 597
Db 121 CAAACAAAGGAATTTTGGGCCACCA-TTGGATTTCAGCAGGTCTCTTCTGACTATAC 179
QY 598 TCCTCCAGAAATTTCTCTTTTGAAGTACTACTGGATTTCATTTGATGGATGCTCTA 657
Db 180 TCCTCCAGAAATTTCTCTTTTGAAGTACTACTGGATTTCATTTGATGGATGCTCTA 239
QY 658 CAGCCTCATGATCTACAGCTGGAAGAAATATCTACTGTGCTGTTTCATATATGGTGG 717
Db 240 CAGCCTCATGATCTACAGCTGGAAGAAATATCTACTGTGCTGTTTCATATATGGTGG 299
QY 718 TCCTCAGTGCAGTGGTGAATTAATCGTTTAAAGAGTCAAGTATTTCCGCTTGAATAC 777
Db 300 TCCTCAGTGCAGTGGTGAATTAATCGTTTAAAGAGTCAAGTATTTCCGCTTGAATAC 359
QY 778 CCTAGCCTCTCTAGGTTATGTGGTTGTAGTAGACACACAGGGGATCCTGTCAACCGAGG 837
Db 360 CCTAGCCTCTCTAGGTTATGTGGTTGTAGTAGACACACAGGGGATCCTGTCAACCGAGG 419
QY 838 GCTTTAAATTTGAAGGCGCCTTTAAATATATAAATGGGTCAAAATAGAAAATTCAGATCAGGT 897
Db 420 GCTTTAAATTTGAAGGCGCCTTTAAATATATAAATGGGTCAAAATAGAAAATTCAGATCAGGT 479
```

898	QY	GGAAGGACTCAATATCTAGCTTCGATATGATTTCAATTCATTCAGTTCAGTCTGTGGGCAT	957
899			
900			
901			
902			
903			
904			
905			
906			
907			
908			
909			
910			
911			
912			
913			
914			
915			
916			
917			
918			
919			
920			
921			
922			
923			
924			
925			
926			
927			
928			
929			
930			
931			
932			
933			
934			
935			
936			
937			
938	QY	CCACGGCTGGTCCTATGGAGGATACCTCTCCCTGATGGCATTAATGCAGAGGTCAGATAT	1017
939			
940			
941			
942			
943			
944			
945			
946			
947			
948			
949			
950			
951			
952			
953			
954			
955			
956			
957			
958			
959			
960			
961			
962			
963			
964			
965			
966			
967			
968			
969			
970			
971			
972			
973			
974			
975			
976			
977			
978			
979			
980			
981			
982			
983			
984			
985			
986			
987			
988			
989			
990			
991			
992			
993			
994			
995			
996			
997			
998			
999			
1000			
1001			
1002			
1003			
1004			
1005			
1006			
1007			
1008	QY	CTTCAG-GGTTGCTATTGCTGGGGGCCCCAGTCACCTCTGTG	1056
1009			
1010			
1011			
1012			
1013			
1014			

898	QY	GAAGGAC	CTCA	AATAT	CTAG	CTTC	CGATAT	GATTT	CA	TTC	AGATCG	TG	GGCAT	957
480	Db	GAAGGAC	CTCA	AATAT	CTAG	CTTC	CGATAT	GATTT	CA	TTC	AGATCG	TG	GGCAT	539
958	QY	CAACGG	CTGG	TCC	TATG	AGAG	ATAC	CTCT	CC	CTG	ATGG	CA	TTC	AGATCG
540	Db	CAACGG	CTGG	TCC	TATG	AGAG	ATAC	CTCT	CC	CTG	ATGG	CA	TTC	AGATCG
1018	QY	CTTCAG	-GGT	TGCT	ATT	GCTGG	GGCC	CCAG	TC	ACT	CT	GTG	1056	
600	Db	CTTCAG	CGGT	TGCT	ATT	GCTGG	GGCC	CCAG	TC	ACT	CT	GTG	639	

480 GGAAGGACTCCAATATCTAGCTTCTCGATATGATTTCATTGACTTAGATCGTGTGGCAT 539

958 CCA CGGCTGGTCCCTATGGAGGATACCTCTCCCTGATGGCATTAA TGCAGAGGTCAGATAT 1017

540 CCACGGCTGTCCTATGGAGGATACCTCTCCCTGATGGCA^{TT}AATGCAGAGGTCAGATAT 599

018 CTTTCAG-GGTTGCTATTGCTGGGGCCCCAGTCACTCTGTG 1056

600 CTTCAGCGGTGCTATTGCTGGGCCCCAGTCACTCTGTG 639

Search completed: May 4, 2006, 07:56:22
Job time : 4060.93 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 3, 2006, 01:01:26 ; Search time 342.808 Seconds
(without alignments)
5615.673 Million cell updates/sec

Title: us-10-825-632-8
Perfect score: 1083
Sequence: 1 ggaagaagatgccagatcag.....tatgatacagatacacgga 1083

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /cgn2_6/ptodata/1/ina/1 COMB.seq:*
- 2: /cgn2_6/ptodata/1/ina/5 COMB.seq:*
- 3: /cgn2_6/ptodata/1/ina/6A COMB.seq:*
- 4: /cgn2_6/ptodata/1/ina/6B COMB.seq:*
- 5: /cgn2_6/ptodata/1/ina/H COMB.seq:*
- 6: /cgn2_6/ptodata/1/ina/PTUS COMB.seq:*
- 7: /cgn2_6/ptodata/1/ina/PP COMB.seq:*
- 8: /cgn2_6/ptodata/1/ina/RE COMB.seq:*
- 9: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1083	100.0	1083	3	US-10-070-464-8
2	1055.8	97.5	4309	3	US-09-976-674-14
3	791.2	73.1	2797	3	US-09-976-594-1103
4	791.2	73.1	3120	3	US-10-070-464-2
5	789.6	72.9	2671	3	US-09-976-674-2
6	779.2	71.9	4829	3	US-09-976-674-12
7	497.2	45.9	4685	3	US-09-976-674-22
8	479.2	44.2	4676	3	US-09-976-674-20
9	459	42.4	1197	3	US-10-070-464-4
10	434.2	40.1	1669	3	US-10-070-464-6
11	428.2	39.5	4523	3	US-09-976-674-8
12	344.2	31.8	2617	3	US-09-976-674-4
13	344.2	31.8	4219	3	US-09-976-674-28
14	344.2	31.8	4302	3	US-09-976-674-24
15	329.8	30.5	4180	3	US-09-976-674-36
16	329.8	30.5	4263	3	US-09-976-674-34
17	311.8	28.8	4076	3	US-09-976-674-32
18	311.8	28.8	4159	3	US-09-976-674-30
19	297.4	27.5	4037	3	US-09-976-674-40
20	297.4	27.5	4120	3	US-09-976-674-38
21	198.4	18.3	823	3	US-09-280-116-171
C 22	131	12.1	480	3	US-09-270-767-684
C 23	131	12.1	480	3	US-09-270-767-15966
C 24	129.2	11.9	612	3	US-09-392-184-31

25	91.4	8.4	2411	3	US-09-976-674-26	Sequence 26, Appl
26	62.6	5.8	2946	3	US-09-902-540-9675	Sequence 9675, Ap
C 27	62.6	5.8	14555	3	US-09-902-540-1096	Sequence 1096, Ap
28	61.2	5.7	2199	3	US-09-518-550-38	Sequence 38, Appl
29	59.6	5.5	3407	3	US-10-002-593-5	Sequence 5, Appli
30	59.6	5.5	3407	3	US-09-949-016-275	Sequence 275, App
31	59.6	5.5	3407	3	US-10-423-714-5	Sequence 5, Appli
32	59.6	5.5	3419	3	US-09-949-016-4579	Sequence 4579, Ap
C 33	58.2	5.4	657	3	US-09-221-017B-646	Sequence 646, App
34	58	5.4	2924	3	US-09-023-655-1026	Sequence 1026, Ap
35	58	5.4	2924	6	PCT-US93-07923-1	Sequence 1, Appli
36	57.8	5.3	2214	3	US-09-902-540-4124	Sequence 4124, Ap
C 37	57.8	5.3	23091	3	US-09-902-540-1204	Sequence 1204, Ap
38	55	5.1	1023	3	US-09-270-767-13726	Sequence 13726, A
39	52.4	4.8	535	3	US-09-389-681-428	Sequence 428, App
40	52.4	4.8	535	3	US-09-620-405B-428	Sequence 428, App
41	52.4	4.8	535	3	US-09-433-826B-428	Sequence 428, App
42	52.4	4.8	535	3	US-09-604-287A-428	Sequence 428, App
43	52.4	4.8	535	3	US-09-834-759-428	Sequence 428, App
44	52.4	4.8	535	3	US-09-590-751A-428	Sequence 428, App
45	52.4	4.8	535	3	US-09-551-621-428	Sequence 428, App

ALIGNMENTS

RESULT 1
US-10-070-464-8
; Sequence 8, Application US/10070464
; Patent No. 6881564
; GENERAL INFORMATION:
; APPLICANT: ABBOTT, Catherine Anne
; APPLICANT: GORRELL, Mark Douglas
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES
; FILE REFERENCE: GH-007
; CURRENT APPLICATION NUMBER: US/10/070,464
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: PCT/AU00/01085
; PRIOR FILING DATE: 2000-09-11
; PRIOR APPLICATION NUMBER: AU PQ5709
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: AU PQ2762
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 1083
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-10-070-464-8

Query Match 100.0%; Score 1083; DB 3; Length 1083;
Best Local Similarity 100.0%; Pred. No. 8.9e-310;
Matches 1083; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAAGAAGATGCCAGATCAGCTGAGTCCGCTACCTTTGTTCTCCAGAGAAATTTGATAG 60
DB 1 GGAAGAAGATGCCAGATCAGCTGAGTCCGCTACCTTTGTTCTCCAGAGAAATTTGATAG 60

QY 61 ATATTCGGCTATTGGTGGTGTCCTCAAAAGCTGAAACAACTCCAGTGGTGGTAAATTTCT 120
DB 61 ATATTCGGCTATTGGTGGTGTCCTCAAAAGCTGAAACAACTCCAGTGGTGGTAAATTTCT 120

QY 121 TAGAATTCATATGAGAAATGATGAATCTGAGGTGGAAATTTATTCATGTTACATCCCC 180
DB 121 TAGAATTCATATGAGAAATGATGAATCTGAGGTGGAAATTTATTCATGTTACATCCCC 180

QY 181 TATGTTGGAACAACAGGAGGGCAGATTCATTCCGTTATCTAAACAGGTGACGAAATCC 240
DB 181 TATGTTGGAACAACAGGAGGGCAGATTCATTCCGTTATCTAAACAGGTGACGAAATCC 240

QY 241 TAAAGTCATCTTTTAAGATGTCAGAAATATGATGATGCTGAAGGAGGATCATGATTGA 300
DB 241 TAAAGTCATCTTTTAAGATGTCAGAAATATGATGATGCTGAAGGAGGATCATGATTGA 300

QY 899 GAAGGACTCCAAATCTAGCTTCTCGATATGATTTCAITGACTTAGATCGTGTGGGCATC 958
DB 1841 GAAGGACTCCAAATCTAGCTTCTCGATATGATTTCAITGACTTAGATCGTGTGGGCATC 1900
QY 959 CACGGCTGGTCTTATGGAGGATACCTCTCCCTGATGGCAATTAATGACAGAGGTCAGATATC 1018
DB 1901 CACGGCTGGTCTTATGGAGGATACCTCTCCCTGATGGCAATTAATGACAGAGGTCAGATATC 1960
QY 1019 TTCAGGGTTGCTATTGCTGGGGCCCAAGTCACCTCTCTGAGTCTTCTATGATACAGGATAC 1078
DB 1961 TTCAGGGTTGCTATTGCTGGGGCCCAAGTCACCTCTCTGAGTCTTCTATGATACAGGATAC 2020
QY 1079 ACGGA 1083
DB 2021 ACGGA 2025

RESULT 3

US-09-976-594-1103
; Sequence 1103, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
; APPLICANT: Furness, Michael
; APPLICANT: Buchbinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/09/976,594
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,409
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 1103
; LENGTH: 2797
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6673549 977951.1
US-09-976-594-1103

Query Match 73.1%; Score 791.2; DB 3; Length 2797;
Best Local Similarity 99.0%; Pred. No. 2.4e-223;
Matches 796; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 280 TGAAGGAGGATCATAGTTCATGAGTCAGAGGCTGGTATATTTGAAGGCACCAAGA 339
DB 1411 TGGATCTTAATATCCAAGTTGATGAAGTCAGAGGCTGGTATATTTGAAGGCACCAAGA 1470
QY 340 CTCGCCCTTTAGAGCATCACCTGTACGTAGTCAGTTACGTAAATCTCGAGAGGTGACAAG 399
DB 1471 CTCGCCCTTTAGAGCATCACCTGTACGTAGTCAGTTACGTAAATCTCGAGAGGTGACAAG 1530
QY 400 GCTGACTGACCGTGGCTACTACATTTCTGCTGCATCAGTACAGCACTGTGACTTCTTTAT 459
DB 1531 GCTGACTGACCGTGGCTACTACATTTCTGCTGCATCAGTACAGCACTGTGACTTCTTTAT 1590
QY 460 AAGTAAGTATAGTAACCAAGAGAAATCCACACTGTGTCCCTTTTCAAGCTATCAAGTCC 519
DB 1591 AAGTAAGTATAGTAACCAAGAGAAATCCACACTGTGTCCCTTTTCAAGCTATCAAGTCC 1650
QY 520 TGAAGATGACCCCAACTTGGCAAAACAAAGGAATTTTGGGCCACCATTTTGGATTTCAGCAGG 579
DB 1651 TGAAGATGACCCCAACTTGGCAAAACAAAGGAATTTTGGGCCACCATTTTGGATTTCAGCAGG 1710
QY 580 TCCTCTTCCTGACTATACCTCTCAGAAATTTCTCTTTTGAAGTACTACTGGAATTTAC 639
DB 1711 TCCTCTTCCTGACTATACCTCTCAGAAATTTCTCTTTTGAAGTACTACTGGAATTTAC 1770
QY 640 ATTCTATGGGATGCTCTTACAGGCTCATGATCTACAGCCTGGAAAGAAATATCTTACTGT 699
DB 1771 ATTCTATGGGATGCTCTTACAGGCTCATGATCTACAGCCTGGAAAGAAATATCTTACTGT 1830

QY 700 GCTGTTTCATATATGCTGCTCCTCAGGTGCGAGTTGGTGAATAAATCGGTTTAAAGGAGTCAA 759
DB 1831 GCTGTTTCATATATGCTGCTCCTCAGGTGCGAGTTGGTGAATAAATCGGTTTAAAGGAGTCAA 1890
QY 760 GTATTTCGCTTGAATACCCCTAGCTCTCTAGGTTATGTTGGTTGTAGTGATAGACAAACAG 819
DB 1891 GTATTTCGCTTGAATACCCCTAGCTCTCTAGGTTATGTTGGTTGTAGTGATAGACAAACAG 1950
QY 820 GGGATCCTGTCCACCGAGGCTTAAATTTGAAGGCGCTTTAAATATATAAATATGCTCAAAAT 879
DB 1951 GGGATCCTGTCCACCGAGGCTTAAATTTGAAGGCGCTTTAAATATATAAATATGCTCAAAAT 2010
QY 880 AGAAATTCAGCATCAGGTGGAAGGACTCCCAATATCTAGCTTCTCGATATGATTTCAATGA 939
DB 2011 AGAAATTCAGCATCAGGTGGAAGGACTCCCAATATCTAGCTTCTCGATATGATTTCAATGA 2070
QY 940 CTTAGATCGTGGGCATCCACGGCTGGTCTCTATGAGGATACCTCTCCCTGATGGCAT 999
DB 2071 CTTAGATCGTGGGCATCCACGGCTGGTCTCTATGAGGATACCTCTCCCTGATGGCAT 2130
QY 1000 AATGCAGAGGTCAGATATCTTCAAGGTTGCTATTGCTGGGGCCCCAGTCACTCTGTGGAT 1059
DB 2131 AATGCAGAGGTCAGATATCTTCAAGGTTGCTATTGCTGGGGCCCCAGTCACTCTGTGGAT 2190
QY 1060 CTTCTATGATACAGGATACACGGA 1083
DB 2191 CTTCTATGATACAGGATACACGGA 2214

RESULT 4

US-10-070-464-2
; Sequence 2, Application US/10070464
; Patent No. 6881564
; GENERAL INFORMATION:
; APPLICANT: ABBOTT, Catherine Anne
; APPLICANT: CORRELL, Mark Douglas
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES
; FILE REFERENCE: GH-007
; CURRENT APPLICATION NUMBER: US/10/070,464
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: PCT/AU00/01085
; PRIOR FILING DATE: 2000-09-11
; PRIOR APPLICATION NUMBER: AU PQ5709
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: AU PQ2762
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 3120
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-10-070-464-2

Query Match 73.1%; Score 791.2; DB 3; Length 3120;
Best Local Similarity 99.0%; Pred. No. 2.5e-223;
Matches 796; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 280 TGAAGGAGGATCATAGTTCATGAGTCAGAGGCTGGTATATTTGAAGGCACCAAGA 339
DB 1740 TGGATCTTAATATCCAAGTTGATGAAGTCAGAGGCTGGTATATTTGAAGGCACCAAGA 1799
QY 340 CTCGCCCTTTAGAGCATCACCTGTACGTAGTCAGTTACGTAAATCTCGAGAGGTGACAAG 399
DB 1800 CTCGCCCTTTAGAGCATCACCTGTACGTAGTCAGTTACGTAAATCTCGAGAGGTGACAAG 1859
QY 400 GCTGACTGACCGTGGCTACTACATTTCTGCTGCATCAGTACAGCACTGTGACTTCTTTAT 459
DB 1860 GCTGACTGACCGTGGCTACTACATTTCTGCTGCATCAGTACAGCACTGTGACTTCTTTAT 1919
QY 460 AAGTAAGTATAGTAACCAAGAGAAATCCACACTGTGTCCCTTTTCAAGCTATCAAGTCC 519
DB 1920 AAGTAAGTATAGTAACCAAGAGAAATCCACACTGTGTCCCTTTTCAAGCTATCAAGTCC 1979

Qy	520	TGAAGATGACCCAACTTGCAAAA	CAAAAGGAATTTTGGCCACCATTTTGGATTACAGCAGG	579
Db	1980	TGAAGATGACCCAACTTGCAAAA	CAAAAGGAATTTTGGCCACCATTTTGGATTACAGCAGG	2039
Qy	580	TCCTCTCTCGTACTACTCTCTCAGAAA	TTTTCTCTTTGGAAGTACTACTGGATTTAC	639
Db	2040	TCCTCTCTCGTACTACTCTCTCAGAAA	TTTTCTCTTTGGAAGTACTACTGGATTTAC	2099
Qy	640	ATTGATGGAGTCCTCTCAAGACCTCA	TGATCTACAGCCTGGAAAGAAATATCTCTACTGT	699
Db	2100	ATTGATGGAGTCCTCTCAAGACCTCA	TGATCTACAGCCTGGAAAGAAATATCTCTACTGT	2159
Qy	700	GCTGTTTCATATATGTGTGCTCTCAGG	TGCAGTTGGTGAATAATTCGGTTTAAAGGAGTCAA	759
Db	2160	GCTGTTTCATATATGTGTGCTCTCAGG	TGCAGTTGGTGAATAATTCGGTTTAAAGGAGTCAA	2219
Qy	760	GTATTTTCGGTTCGAAATACCCTAGC	CTCTAGGTTATGTGGTTAGTAGTAGACAACAG	819
Db	2220	GTATTTTCGGTTCGAAATACCCTAGC	CTCTAGGTTATGTGGTTAGTAGTAGACAACAG	2279
Qy	820	GGGATCCTGTCAACCGAGGGCTTAAAT	TTTGAAGCGCGCTTTTAAATATAAAATGGGTCAAAT	879
Db	2280	GGGATCCTGTCAACCGAGGGCTTAAAT	TTTGAAGCGCGCTTTTAAATATAAAATGGGTCAAAT	2339
Qy	880	AGAAATTGACGATCAGGTGGAAGAC	TCCCAATATCTAGCTTCCGATATGATTTTCATGA	939
Db	2340	AGAAATTGACGATCAGGTGGAAGAC	TCCCAATATCTAGCTTCCGATATGATTTTCATGA	2399
Qy	940	CTTAGATCGTGTGGGCATCCAGCGCT	GGTCCCTATGAGGATACCTCTCCCTGATGGCAATT	999
Db	2400	CTTAGATCGTGTGGGCATCCAGCGCT	GGTCCCTATGAGGATACCTCTCCCTGATGGCAATT	2459
Qy	1000	AATCAGAGGTCAGATATCTTTCAGGG	TGCTATTGCTGGGGCCCCAGTCACTCTGTGGAT	1059
Db	2460	AATCAGAGGTCAGATATCTTTCAGGG	TGCTATTGCTGGGGCCCCAGTCACTCTGTGGAT	2519
Qy	1060	CTTCTATGATACAGGATACACGA	1083	
Db	2520	CTTCTATGATACAGGATACACGA	2543	

RESULT 5

```

US-09-976-674-2
/ Sequence 2, Application US/09976674
/ Patent No. 6844180
/ GENERAL INFORMATION:
/ APPLICANT: Qi, Steve
/ APPLICANT: Akinsanya, Karen
/ APPLICANT: Riviere, Pierre
/ APPLICANT: Junien, Jean-Louis
/ TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
/ FILE REFERENCE: 70669
/ CURRENT APPLICATION NUMBER: US/09/976,674
/ CURRENT FILING DATE: 2001-10-12
/ ERROR APPLICATION NUMBER: US 60/240,117
/ ERROR FILING DATE: 2000-10-12
/ PRIOR SEQ ID NOS: 61
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO. 2
/ LENGTH: 2671
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-976-674-2

```

	Query Match	72.9%	Score 789.6;	DB 3;	Length 2671;
	Best Local Similarity	98.9%;	Pred. No. 6.9e-223;		
	Matches 795;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	280	TTAAGGAGGATCATAGTTGAAGTCAGAAGCGCTGGTATATTTTGAAGGCACCAAGA	339		
Db	1534	TGGATCTAATATCCAAGTGTGTAAGTCAGAAGCGCTGGTATATTTTGAAGGCACCAAGA	1593		

QY	340	CTCCCTTTTAGAGCATCACTGTACGTGACGTACGTAAATCTCGAGAGGTGACAAG	399
DB	1594	CTCCCTTTTAGAGCATCACTGTACGTGACGTACGTAAATCTCGAGAGGTGACAAG	1653
QY	400	GCTGACTGACCGTGGCTACTACACATCTTTGCTGTCATCAGTCAGCACTCTGTACTCTTTTAT	459
DB	1654	GCTGACTGACCGTGGCTACTACACATCTTTGCTGTCATCAGTCAGCACTCTGTACTCTTTTAT	1713
QY	460	AAGTAAGTATAGTAAACAGAGAAGATCCACACTGTGTGTCCTTTTAAAGACTATCAAGTCC	519
DB	1714	AAGTAAGTATAGTAAACAGAGAAGATCCACACTGTGTGTCCTTTTAAAGACTATCAAGTCC	1773
QY	520	TGAAGTAGACCAACTGTGCAAAACAAAGGAATTTTGGGCCACCATTTTGGATTCCAGCAGG	579
DB	1774	TGAAGTAGACCAACTGTGCAAAACAAAGGAATTTTGGGCCACCATTTTGGATTCCAGCAGG	1833
QY	580	TCTCTTCTGACTATACTCTCCAGAAATTTTCTTTTGAAGTACTACTCGATTATAC	639
DB	1834	TCTCTTCTGACTATACTCTCCAGAAATTTTCTTTTGAAGTACTACTCGATTATAC	1893
QY	640	ATTGTATGGGATGCTCTACAGGCTCATGATCTACAGCTCGGAAGAATAATCCTACTGT	699
DB	1894	ATTGTATGGGATGCTCTACAGGCTCATGATCTACAGCTCGGAAGAATAATCCTACTGT	1953
QY	700	GCTGTTCATATATGCTGGTCTCAGTGCAGTTTGTGTAATAATCGTTTAAAGAGTCAA	759
DB	1954	GCTGTTCATATATGCTGGTCTCAGTGCAGTTTGTGTAATAATCGTTTAAAGAGTCAA	2013
QY	760	GTATTTCCGCTTGAAATACCTACGCTCTCTAGGTTATGTGTTGTAGTGATAGACAACAG	819
DB	2014	GTATTTCCGCTTGAAATACCTACGCTCTCTAGGTTATGTGTTGTAGTGATAGACAACAG	2073
QY	820	GGGATCTCTGTCCACGAGGGCTTAAATTTGAAGGCGCCTTTAAATATATAAATGGGTCAAAT	879
DB	2074	GGGATCTCTGTCCACGAGGGCTTAAATTTGAAGGCGCCTTTAAATATATAAATGGGTCAAAT	2133
QY	880	AGAAATTCACCATCAGGTGGAAGGATCTCAAATATCTAGCTTCTCGATATGATTTCAATTGA	939
DB	2134	AGAAATTCACCATCAGGTGGAAGGATCTCAAATATCTAGCTTCTCGATATGATTTCAATTGA	2193
QY	940	CTTAGATCGTGTGGGCATCCAGCGCTGGTCTTATGGAGGATACCTCTCCCTGATGGCAAT	999
DB	2194	CTTAGATCGTGTGGGCATCCAGCGCTGGTCTTATGGAGGATACCTCTCCCTGATGGCAAT	2253
QY	1000	AATGCAGAGGTTCAGATATCTTCAGGGTTGCTATTCTCGGGGCCACGCTACTCTGTGGAT	1059
DB	2254	AATGCAGAGGTTCAGATATCTTCAGGGTTGCTATTCTCGGGGCCACGCTACTCTGTGGAT	2313
QY	1060	CTTCTATGATACAGGATACACGGA	1083
DB	2314	CTTCTATGATACAGGATACACGGA	2337
RESULT 6			
US-09-976-674-12			
; Sequence 12, Application US/09976674			
; Patent No. 684180			
; GENERAL INFORMATION:			
; APPLICANT: Qi, Steve			
; APPLICANT: Akinsanya, Karen			
; APPLICANT: Riviere, Pierre			
; APPLICANT: Junien, Jean-Louis			
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV			
; FILE REFERENCE: 70669			
; CURRENT APPLICATION NUMBER: US/09/976,674			
; CURRENT FILING DATE: 2001-10-12			
; PRIOR APPLICATION NUMBER: US 60/240,117			
; PRIOR FILING DATE: 2000-10-12			
; NUMBER OF SEQ ID NOS: 61			
; SOFTWARE: PatentIn version 3.1			
; SEQ ID NO 12			
; LENGTH: 4829			
; TYPE: DNA			

RESULT 6

```

US-09-976-674-12
; Sequence 12, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPP1V
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 4829
; TYPE: DNA

```

```

; ORGANISM: Homo sapiens
US-09-976-674-12

Query Match      71.9%; Score 779.2; DB 3; Length 4829;
Best Local Similarity 98.8%; Pred. No. 1e-219;
Matches 796; Conservative 0; Mismatches 8; Indels 2; Gaps 1;

Qy 280 TGAAGGAAGGATCATATGTTGATGAAGTCAGAAAGGCTGGTATATTTTGAAGGCCACCAAGA 339
Db 1740 TGGATCTTAATATCAAGTTTGATGAAGTCAGAAAGGCTGGTATATTTTGAAGGCCACCAAGA 1799

Qy 340 CTCCCTTTAGAGCATCACCTGTACGTACGTACGTACGTACGTACGTACGTACGTACGTACGT 399
Db 1800 CTCCCTTTAGAGCATCACCTGTACGTACGTACGTACGTACGTACGTACGTACGTACGTACGT 1859

Qy 400 GCTGACTGACCGTGGCTACTACATTTCTGCTGCTACGTACGTACGTACGTACGTACGTACGT 459
Db 1860 GCTGACTGACCGTGGCTACTACATTTCTGCTGCTACGTACGTACGTACGTACGTACGTACGT 1919

Qy 460 AAGTAAGTATAGTAACACAGAAATCCACACTGTGTGTCCTTTTACAGCTATCAAGTCC 519
Db 1920 AAGTAAGTATAGTAACACAGAAATCCACACTGTGTGTCCTTTTACAGCTATCAAGTCC 1979

Qy 520 TGAAGTACCCCAACTTGCACAAAGAAATTTTGGGCCACCAATTTTGGATTTCAGCAGG 579
Db 1980 TGAAGTACCCCAACTTGCACAAAGAAATTTTGGGCCACCAATTTTGGATTTCAGCAGG 2039

Qy 580 TCCTCTTCCTGACTACTACTCTCCAGAAATTTTCTCTTTGAAAGTACTACTGGAATTTAC 639
Db 2040 TCCTCTTCCTGACTACTACTCTCCAGAAATTTTCTCTTTGAAAGTACTACTGGAATTTAC 2099

Qy 640 ATTGTATGGGATGCTCTTACAGCCTCATGATCTACAGCCTGGAAGAAATATCCTACTGT 699
Db 2100 ATTGTATGGGATGCTCTTACAGCCTCATGATCTACAGCCTGGAAGAAATATCCTACTGT 2159

Qy 700 GCTGTTTCATATATGGTGG--TCCTCAGGTGCAATTTGGTGAATATCGTTTAAAGGATC 757
Db 2160 GCTGTTTCATATATGGTGGTCTCCTCAGGTGCAATTTGGTGAATATCGTTTAAAGGATC 2219

Qy 758 AAGTATTTCCGCTTGAATACCTTAGCTCTCTAGGTATGTTGTTAGTATAGATAGACAAC 817
Db 2220 AAGTATTTCCGCTTGAATACCTTAGCTCTCTAGGTATGTTGTTAGTATAGATAGACAAC 2279

Qy 818 AGGGATCTCTGTACCGAGGGCTTAAATTTGAAGCGGCTTTAAATATAAAATGGGTCAA 877
Db 2280 AGGGATCTCTGTACCGAGGGCTTAAATTTGAAGCGGCTTTAAATATAAAATGGGTCAA 2339

Qy 878 ATAGAAATTCAGCATCAGGTGGAAGCACTCCAATATCTAGCTTCTCGATATGATTTCAAT 937
Db 2340 ATAGAAATTCAGCATCAGGTGGAAGCACTCCAATATCTAGCTTCTCGATATGATTTCAAT 2399

Qy 938 GACTTAGATCGTGTGGGCATCCACGGCTGCTCTATGGAGGATACCTCTCCCTGATGGCA 997
Db 2400 GACTTAGATCGTGTGGGCATCCACGGCTGCTCTATGGAGGATACCTCTCCCTGATGGCA 2459

Qy 998 TTAATGAGAGGTGAGATATCTTTCAGGGTGTCTATGCTGGGGCCCCAGTCACTCTGTGG 1057
Db 2460 TTAATGAGAGGTGAGATATCTTTCAGGGTGTCTATGCTGGGGCCCCAGTCACTCTGTGG 2519

Qy 1058 ATCTTCTATGATACAGGATACACGGA 1083
Db 2520 ATCTTCTATGATACAGGATACACGGA 2545
```

```

RESULT 7
US-09-976-674-22
; Sequence 22, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Oi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
```

```

; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 22
; LENGTH: 4685
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-976-674-22

Query Match      45.9%; Score 497.2; DB 3; Length 4685;
Best Local Similarity 81.3%; Pred. No. 2.3e-136;
Matches 654; Conservative 0; Mismatches 8; Indels 142; Gaps 1;

Qy 280 TGAAGGAAGGATCATATGTTGATGAAGTCAGAAAGGCTGGTATATTTTGAAGGCCACCAAGA 339
Db 1740 TGGATCTTAATATCAAGTTTGATGAAGTCAGAAAGGCTGGTATATTTTGAAGGCCACCAAGA 1799

Qy 340 CTCCCTTTAGAGCATCACCTGTACGTACGTACGTACGTACGTACGTACGTACGTACGTACGT 399
Db 1800 CTCCCTTTAGAGCATCACCTGTACGTACGTACGTACGTACGTACGTACGTACGTACGTACGT 1859

Qy 400 GCTGACTGACCGTGGCTACTACATTTCTGCTGCTACGTACGTACGTACGTACGTACGTACGT 459
Db 1860 GCTGACTGACCGTGGCTACTACATTTCTGCTGCTACGTACGTACGTACGTACGTACGTACGT 1919

Qy 460 AAGTAAGTATAGTAACACAGAAATCCACACTGTGTGTCCTTTTACAGCTATCAAGTCC 519
Db 1920 AAGTAAGTATAGTAACACAGAAATCCACACTGTGTGTCCTTTTACAGCTATCAAGTCC 1979

Qy 520 TGAAGTACCCCAACTTGCACAAAGAAATTTTGGGCCACCAATTTTGGATTTCAGCAGG 579
Db 1980 TGAAGTACCCCAACTTGCACAAAGAAATTTTGGGCCACCAATTTTGGATTTCAGCAGG 2034

Qy 580 TCCTCTTCCTGACTACTACTCTCCAGAAATTTTCTCTTTGAAAGTACTACTGGAATTTAC 639
Db 2035 TCCTCTTCCTGACTACTACTCTCCAGAAATTTTCTCTTTGAAAGTACTACTGGAATTTAC 2034

Qy 640 ATTGTATGGGATGCTCTTACAGCCTCATGATCTACAGCCTGGAAGAAATATCCTACTGT 699
Db 2035 ATTGTATGGGATGCTCTTACAGCCTCATGATCTACAGCCTGGAAGAAATATCCTACTGT 2034

Qy 700 GCTGTTTCATATATGGTGGTCTCCTCAGGTGCAATTTGGTGAATATCGTTTAAAGGATCAA 759
Db 2035 GCTGTTTCATATATGGTGGTCTCCTCAGGTGCAATTTGGTGAATATCGTTTAAAGGATCAA 2077

Qy 760 GTATTTCCGCTTGAATACCTTAGCTCTCTAGGTATGTTGTTAGTATAGATAGACAACAG 819
Db 2078 GTATTTCCGCTTGAATACCTTAGCTCTCTAGGTATGTTGTTAGTATAGATAGACAACAG 2137

Qy 820 GGGATCTCTGTACCGAGGGCTTAAATTTGAAGCGGCTTTAAATATAAAATGGGTCAAAT 879
Db 2138 GGGATCTCTGTACCGAGGGCTTAAATTTGAAGCGGCTTTAAATATAAAATGGGTCAAAT 2197

Qy 880 AGAAATTGACGATCAGGTGGAAGGACTCCCAATATCTAGCTTCTCGATATGATTTCAATGA 939
Db 2198 AGAAATTGACGATCAGGTGGAAGGACTCCCAATATCTAGCTTCTCGATATGATTTCAATGA 2257

Qy 940 CTTAGATCGTGTGGGCATCCACGGCTGCTCTATGAGGATACCTCTCCCTGATGGCATTT 999
Db 2258 CTTAGATCGTGTGGGCATCCACGGCTGCTCTATGAGGATACCTCTCCCTGATGGCATTT 2317

Qy 1000 AATGACAGAGGTGAGATATCTTTCAGGGTGTCTATGCTGGGGCCCCAGTCACTCTGTGGAT 1059
Db 2318 AATGACAGAGGTGAGATATCTTTCAGGGTGTCTATGCTGGGGCCCCAGTCACTCTGTGGAT 2377

Qy 1060 CTTCTATGATACAGGATACACGGA 1083
Db 2378 CTTCTATGATACAGGATACACGGA 2401
```

```
RESULT 8
US-09-976-674-20
; Sequence 20, Application US/09976674
; Patent No. 684180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20
; LENGTH: 4676
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-976-674-20

Query Match      44.2%; Score 479.2; DB 3; Length 4676;
Best Local Similarity 80.2%; Pred. No. 4.9e-131;
Matches 645; Conservative 0; Mismatches 8; Indels 151; Gaps 1;

QY 280 TGAAGGAAGGATCATAGTTGATGAAGTCAAGAGGCTGGTATATTTTGAAGGCCCAAGA 339
DB 1740 TGGATCTAATATCAAGTTGATGAAGTCAAGAGGCTGGTATATTTTGAAGGCCCAAGA 1799
QY 340 CTCCTCTTTAGAGCATCACCCTGCTAGTGTACCTTAATCTCTGAGAGGTGACAG 399
DB 1800 CTCCTCTTTAGAGCATCACCCTGCTAGTGTACCTTAATCTCTGAGAGGTGACAG 1859
QY 400 GCTGACTGACCTGGCTATCATCTTCTGCTGATCATGCTGAGCACTGTGACTCTTTAT 459
DB 1860 GCTGACTGACCTGGCTATCATCTTCTGCTGATCATGCTGAGCACTGTGACTCTTTAT 1919
QY 460 AGTAAAGTATAGTAACAGAGAAATCCACATCTGTGTGCTCTTTTGAAGTATCAAGTCC 519
DB 1920 AGTAAAGTATAGTAACAGAGAAATCCACATCTGTGTGCTCTTTTGAAGTATCAAGTCC 1979
QY 520 TGAAGATGACCCAACTTGCAAAACAAAGAAATTTTGGCCCAACCAATTTGGATTACAGG 579
DB 1980 TGAAGATGACCCAACTTGCAAAACAAAGAAATTTTGGCCCAACCAATTTGGATTACAGG 2039
QY 580 TCCTCTTCTGACTATCTCTCAAGCCTCATGATCTACAGCCTGCAAGAAATTTTGAAGTATCA 639
DB 2040 TCCTCTTCTGACTATCTCTCAAGCCTCATGATCTACAGCCTGCAAGAAATTTTGAAGTATCA 2099
QY 640 ATTGTATGGGATGCTCTACAAGCCTCATGATCTACAGCCTGCAAGAAATTTTGAAGTATCA 699
DB 2100 ATTGTATGGGATGCTCTACAAGCCTCATGATCTACAGCCTGCAAGAAATTTTGAAGTATCA 2159
QY 700 GCTGTTTCATATATGGTGGTCTCAGTGTGCTGATTAATCGGTTTAAAGAGTCAA 759
DB 2160 GCTGTTTCATATATGGTGGTCTCAGTGTGCTGATTAATCGGTTTAAAGAGTCAA 2179
QY 760 GTATTTCCGCTTGAATACCTAGCCTCTCTAGTTATGTGTGTTAGTATAGACAAACAG 819
DB 2180 ----- 2179
QY 820 GGGATCCTGTCAACGAGGCTTAAATTTGAAGGCGCCTTTAAATATATAAATGGTCAAT 879
DB 2180 -----GGGTCAAT 2188
QY 880 AGAAATGAGATCAGGTGGAGGACTCCAAATATCTAGCTTCTCGATATGATTTCAATGA 939
DB 2189 AGAAATGAGATCAGGTGGAGGACTCCAAATATCTAGCTTCTCGATATGATTTCAATGA 2248
```

```
QY 940 CTTAGATCCTGTGGGCATCCACGGCTGCTCTATAGAGAGATACCTCTCCCTGATGCATT 999
DB 2249 CTTAGATCCTGTGGGCATCCACGGCTGCTCTATAGAGAGATACCTCTCCCTGATGCATT 2308
QY 1000 AATGCAGAGTCAAGATATCTTCAGAGGTGCTATTGCTGGGGCCCCAGTCACTCTGTGGAT 1059
DB 2309 AATGCAGAGTCAAGATATCTTCAGAGGTGCTATTGCTGGGGCCCCAGTCACTCTGTGGAT 2368
QY 1060 CTTCTATGATACAGGATACACGGA 1083
DB 2369 CTTCTATGATACAGGATACACGGA 2392

RESULT 9
US-10-070-464-4
; Sequence 4, Application US/10070464
; Patent No. 6881564
; GENERAL INFORMATION:
; APPLICANT: ABBOTT, Catherine Anne
; APPLICANT: GORRELL, Mark Douglas
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES
; FILE REFERENCE: GH-007
; CURRENT APPLICATION NUMBER: US/10/070,464
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: PCT/AU00/01085
; PRIOR FILING DATE: 2000-09-11
; PRIOR APPLICATION NUMBER: AU PQ5709
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: AU PQ2762
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1197
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-10-070-464-4

Query Match      42.4%; Score 459; DB 3; Length 1197;
Best Local Similarity 80.7%; Pred. No. 2.6e-125;
Matches 616; Conservative 0; Mismatches 0; Indels 147; Gaps 1;

QY 321 ATTTTGAAGGCACCAAGAGACTCCCTTTTAGAGCATCACCTGTACGTAGTACGTAA 380
DB 1 ATTTTGAAGGCACCAAGAGACTCCCTTTTAGAGCATCACCTGTACGTAGTACGTAA 60
QY 381 ATCTGTGAGAGGTGACAAAGGCTGACTGACCTGGCTACTCATTCTTTGCTGCATCAGTC 440
DB 61 ATCTGTGAGAGGTGACAAAGGCTGACTGACCTGGCTACTCATTCTTTGCTGCATCAGTC 120
QY 441 AGCACTGTGACTTCTTTTATAAGTAAGTATAGTAACCAAGAAATCCACACTGTGTCCC 500
DB 121 AGCACTGTGACTTCTTTTATAAGTAAGTATAGTAACCAAGAAATCCACACTGTGTCCC 180
QY 501 TTTTACAGCTATCAAGTCTGAGATGACCCAACTTGCACCAAGAAATTTTGGGCCA 560
DB 181 TTTTACAGCTATCAAGTCTGAGATGACCCAACTTGCACCAAGAAATTTTGGGCCA 240
QY 561 CCATTTTGGATTCAGCAGGCTCTCTCTGACTATACCTCCAGAAATTTTCTCTTTTG 620
DB 241 CCATTTTGGATTCAGCAGGCTCTCTCTGACTATACCTCCAGAAATTTTCTCTTTTG 300
QY 621 AAAGTACTACTGATTTTCAATTTGATGGATGCTCTACAAGCCTCATGATCTACAGCTG 680
DB 301 AAAGTACTACTGATTTTCAATTTGATGGATGCTCTACAAGCCTCATGATCTACAGCTG 360
QY 681 GAAAGAAATATCTACTGCTGTGTTCAATATATGTTGGTCTCAGGTGAGTGTGTAATA 740
DB 361 GAAAGAAATATCTACTGCTGTGTTCAATATATGTTGGTCTCAGGTGAGTGTGTAATA 403
QY 741 ATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCCCTAGCCTCTCTAGTTTATGTGG 800
DB 404 ----- 403
```

801 TTGTAGTATAGACAAACAGGGGATCCCTGTCCACCGAGGGCTTAAATTTGAAGCGCCTTTA 860
404 ----- 403
861 ATATATAAATGGTCAAAATAGAAATTCAGCATCAGGTGGAGGACTCCCAATATCTAGCTT 920
404 -----GGTCAATATAGAAATTCAGCATCAGGTGGAGGACTCCCAATATCTAGCTT 453
921 CTCGATATGATTTTCAATTCAGTATAGATCGTGTGGGATCCACGGCTGGTCTTATGGAGAT 980
454 CTCGATATGATTTTCAATTCAGTATAGATCGTGTGGGATCCACGGCTGGTCTTATGGAGAT 513
981 ACTCTCCCTGATGATTAATTCAGAGGTCAGATATCTTCAGGTTTGTATTTGCTGGGG 1040
514 ACTCTCCCTGATGATTAATTCAGAGGTCAGATATCTTCAGGTTTGTATTTGCTGGGG 573
1041 CCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGA 1083
574 CCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGA 616
RESULT 10
US-10-070-464-6
; Sequence 6, Application US/10070464
; Patent No. 6881564
; GENERAL INFORMATION:
; APPLICANT: ABBOTT, Catherine Anne
; APPLICANT: GORRELL, Mark Douglas
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES
; FILE REFERENCE: GH-007
; CURRENT APPLICATION NUMBER: US/10/070,464
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: PCT/AU00/01085
; PRIOR FILING DATE: 2000-09-11
; PRIOR APPLICATION NUMBER: AU PQ5709
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: AU PQ2762
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1669
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-10-070-464-6
Query Match 40.1%; Score 434.2; DB 3; Length 1669;
Best Local Similarity 98.2%; Pred. No. 6.3e-118;
Matches 439; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 280 TGAAGGAAGGATCATAGTTGATGAGTCAAGTCAAGAGGCTGGTATATTTTGAAGGCCACCAAGA 339
DB 577 TGGATCTAATATCCAAAGTTGATGAAGTCAAGAGGCTGGTATATTTTGAAGGCCACCAAGA 636
QY 340 CTCCTCTTTAGAGCATCACCTGTACGTAGTCACTTACGTAAATCTCGAGAGGTGACAAG 399
DB 637 CTCCTCTTTAGAGCATCACCTGTACGTAGTCACTTACGTAAATCTCGAGAGGTGACAAG 696
QY 400 GCTGACTGACCGGTGACTTCAATTTCTGTGCTGCATCAGTCAAGCTGTGACTTCTTTAT 459
DB 697 GCTGACTGACCGGTGACTTCAATTTCTGTGCTGCATCAGTCAAGCTGTGACTTCTTTAT 756
QY 460 AAGTAAGTATAGTAACCAAGAGAAATCCACACTGTGTGCCCTTTTCAAGCTATCAAGTCC 519
DB 757 AAGTAAGTATAGTAACCAAGAGAAATCCACACTGTGTGCCCTTTTCAAGCTATCAAGTCC 816
QY 520 TGAAGATGACCCAACTTGCACCAAGAGAAATTTTGGGCCACCAATTTTGAATTCAGCAGG 579
DB 817 TGAAGATGACCCAACTTGCACCAAGAGAAATTTTGGGCCACCAATTTTGAATTCAGCAGG 876
QY 580 TCCTCTTCTGACTATATCTCTCCAGAAATTTTCTTTTGAAGTACTACTGATTTAC 639

877 TCCTCTTCTGACTATATCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGATTAC 936
640 ATTGTATGGGATGCTCTACAGGCTCTATATCTACAGGCTGGAAAGAAATATCTACTCTGT 699
937 ATTGTATGGGATGCTCTACAGGCTCTATATCTACAGGCTGGAAAGAAATATCTACTCTGT 996
700 GCTGTTTCATATATGTTGGTCTCTCAGGT 726
997 GCTGTTTCATATATGTTGGTCTCTCAGGT 1023
RESULT 11
US-09-976-674-8
; Sequence 8, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 4523
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-976-674-8
Query Match 39.5%; Score 428.2; DB 3; Length 4523;
Best Local Similarity 96.1%; Pred. No. 5.7e-116;
Matches 439; Conservative 0; Mismatches 18; Indels 0; Gaps 0;
QY 280 TGAAGGAAGGATCATAGTTGATGAAGTCAAGAGGCTGGTATATTTTGAAGGCCACCAAGA 339
DB 1740 TGGATCTAATATCCAAAGTTGATGAAGTCAAGAGGCTGGTATATTTTGAAGGCCACCAAGA 1799
QY 340 CTCCTCTTTAGAGCATCACCTGTACGTAGTCACTTACGTAAATCTCGAGAGGTGACAAG 399
DB 1800 CTCCTCTTTAGAGCATCACCTGTACGTAGTCACTTACGTAAATCTCGAGAGGTGACAAG 1859
QY 400 GCTGACTGACCGGTGCTACTCAATTTCTGTCATCAGTCAAGCTGTGACTTCTTTAT 459
DB 1860 GCTGACTGACCGGTGCTACTCAATTTCTGTCATCAGTCAAGCTGTGACTTCTTTAT 1919
QY 460 AAGTAAGTATAGTAACCAAGAGAAATCCACACTGTGTGCCCTTTTCAAGCTATCAAGTCC 519
DB 1920 AAGTAAGTATAGTAACCAAGAGAAATTTTGGGCCACCAATTTTGGATTTCAGCAGG 1979
QY 520 TGAAGATGACCCAACTTGCACCAAGAGAAATTTTGGGCCACCAATTTTGGATTTCAGCAGG 579
DB 1980 TGAAGATGACCCAACTTGCACCAAGAGAAATTTTGGGCCACCAATTTTGGATTTCAGCAGG 2039
QY 580 TCCTCTTCTGACTATATCTCTCCAGAAATTTTCTTTTGAAGTACTACTGATTTAC 639
DB 2040 TCCTCTTCTGACTATATCTCTCCAGAAATTTTCTTTTGAAGTACTACTGATTTAC 2099
QY 640 ATTGTATGGGATGCTCTACAGGCTCTATGATCTACAGGCTGGAAAGAAATATCTACTCTGT 699
DB 2100 ATTGTATGGGATGCTCTACAGGCTCTATGATCTACAGGCTGGAAAGAAATATCTACTCTGT 2159
QY 700 GCTGTTTCATATATGTTGGTCTCTCAGGTGCAATTTGGTG 736
DB 2160 GCTGTTTCATATATGTTGGTCTCTCAGGTGCAATTTGGTG 2196
RESULT 12
US-09-976-674-4

```
; Sequence 4, Application US/09976674
; Patent No.: 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 2617
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-976-674-4

Query Match      31.8%; Score 344.2; DB 3; Length 2617;
Best Local Similarity 64.5%; Pred. No. 38-91;
Matches 514; Conservative 0; Mismatches 283; Indels 0; Gaps 0;

QY 287 AGGATCATAGTTCATGAAGTCAGAGGCTGGTATATTTTGAAGGCACCAAAGACTCCCT 346
DB 1517 AAGATCTGGGTCAATGAGGACCAAGCTGGTGTACTTCCAGGGCACCAAGGACACGCCG 1576
QY 347 TTAGAGCATCACCTGTACGTAGTTCAGTTACGTPAAATCCTGGAGAGGTGACAAGGCTGACT 406
DB 1577 CTGGAGCACCATCTACGTGTGTGCTAGCTATGAGGCGCGCGGAGATCGTACGCTCAC 1636
QY 407 GACCGTGGCTACTCATATTTTGTGTCATCAGTCAGCACTGTGACTTCTTTTATAAGTAAG 466
DB 1637 ACGCCCGGGTCTCCCATAGCTGCTCCATGAGCGCAGAACTTCGACATGTTTCGTACGCCAC 1696
QY 467 TATAGTAACCAAGAAGATCCACACTGTGTGTGCTCCCTTTACAAGCTATCAAGTCTCGAAGAT 526
DB 1697 TACAGCAGCGTGAAGCAGCGCGCTCGTGCACGCTTACAAGCTGAGCGGCCCGAGCAG 1756
QY 527 GACCCAACTTGCAAAACAAGAAATTTTGGGCCCAACATTTTGGATTCAGCAGGTCTCTTT 586
DB 1757 GACCCCTCGACAAGCAGCGCCGCTTCTGGGCTAGCATGATGAGGCGAGCCAGCTGCCCC 1816
QY 587 CCTGACTATACCTCCAGAAATTTCTCTTTGAAAGTACTACTGGATTACATTTGAT 646
DB 1817 CGGATTTATGTTCTCCAGAGATCTTCCATTTCCACAGCGCTCGGATGTGGGCTCTAC 1876
QY 647 GGGATGCTCTACAGCCTCATGATCTACAGCCTGGAAGAAATATCCTACTGTGCTGTTTC 706
DB 1877 GGCATGATCTACAAGCCCAAGCCTTGACGACGAGGAAGACCCACCGTCTCTTT 1936
QY 707 ATATATGGTGTCCTCAGTGCAGTTGGTGAATATCGGTTTAAAGAGTCAAGTATTTC 766
DB 1937 GTATATGGAGGCCCCCAGGTGCAGCTGGTGAATTAATCTTCAAGGGCATCAAGTACTTG 1996
QY 767 CGCTTGAATACCTGACCTCTCTAGGTTATGTGGTTGTAGTGATAGACAACAGGGGATCC 826
DB 1997 CGGCTCAACACACTGGCTCTCCCTGGGCTACGCGTGGTGTGATTCACGCGCAGGGGCTCC 2056
QY 827 TGTCAACGAGGCTTAAATTTGAAGCGCCCTTTAAATATATAAAATGGGTCAAAATAGAAATT 886
DB 2057 TGTCAACGAGGCTTCCGTTTCAAGAGGGGCCCTGAAAAACCAAAATGGGCGAGGTGAGATC 2116
QY 887 GACGATCAGGTGGAGGAGCTCCAATATCTAGCTTCTCGATATGATTTCAATTCATTTAGAT 946
DB 2117 GAGGACCAAGGTGGAGGGCTCGAGTTTCGTGGCGAGAAAGTATGGCTTTCATCGACCTGAGC 2176
QY 947 CGTGTGGGATCCACGGCTGCTCTATGAGGAGTACTCTCTCCCTGATGGCATTTATGAG 1006
DB 2177 CGAGTTGCCATCCATGGCTGCTCTACGGGGCTTCTCTCTCGCTCATGGGCTATATCCAC 2236
```

```
QY 1007 AGGTCAAGATATCTTCAGGTTGCTATTTGCTGGGGCCCAAGTCACTCTGTGGATCTCTAT 1066
DB 2237 AAGCCCCAGGTGTTCAAGGTGGCAATCGCGGGTCCCCCGTCCCGTCTGGATGCCCTAC 2296
QY 1067 GATACAGGATACACGGA 1083
DB 2297 GACACAGGGTACACTGA 2313

RESULT 13
US-09-976-674-28
; Sequence 28, Application US/09976674
; Patent No.: 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 28
; LENGTH: 4219
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-976-674-28

Query Match      31.8%; Score 344.2; DB 3; Length 4219;
Best Local Similarity 64.5%; Pred. No. 3.7e-91;
Matches 514; Conservative 0; Mismatches 283; Indels 0; Gaps 0;

QY 287 AGGATCATAGTTCATGAAGTCAGAGGCTGGTATATTTTGAAGGCACCAAAGACTCCCT 346
DB 1873 AAGATCTGGGTCAATGAGGACCAAGCTGGTGTACTTCCAGGGCACCAAGGACACGCCG 1932
QY 347 TTAGAGCATCACCTGTACGTAGTTCAGTTACGTPAAATCCTGGAGAGGTGACAAGGCTGACT 406
DB 1933 CTGGAGCACCATCTCTAGTGTGTGCTAGCTATGAGGCGCGCGGAGATCGTACGCTCAC 1992
QY 407 GACCGTGGCTACTCATATTTTGTGTCATCAGTCAGCACTGTGACTTCTTTTATAAGTAAG 466
DB 1993 ACGCCCGGGTCTCCCATAGCTGCTCCATGAGCCAGAACTTCGACATGTTTCGTACGCCAC 2052
QY 467 TATAGTAACCAAGAAGATCCACACTGTGTGCTCCCTTTACAAGCTATCAAGTCTCTGAAGAT 526
DB 2053 TACAGCAGCGTGAAGCAGCGCGCTTGGTGCACTCTACAAGCTGAGCGGCCCGAGCAG 2112
QY 527 GACCCAACTTGCAAAACAAGAAATTTTGGGCCCAACATTTTGGATTCAGCAGGTCTCTTT 586
DB 2113 GACCCCTCGACAAGCAGCGCCGCTTCTGGGCTAGCATGATGAGGCGAGCCAGCTGCC 2172
QY 587 CCTGACTATACCTCCAGAAATTTCTCTTTTGAAGTACTACTGGATTACATTTGAT 646
DB 2173 CGGATTTATGTTCTCCAGAGATCTTCCATTTCCACAGCGCTCGGATGTGGGCTCTAC 2232
QY 647 GGGATGCTCTACAAGCCTCATGATCTACGCTGGAAGAAATATCCTACTGTGCTGTTTC 706
DB 2233 GGCATGATCTACAAGCCCCACGCCCTTGACGCCAGGGAAGAACCCACCCCTCTTT 2292
QY 707 ATATATGGTGTCCTCAGTGCAGTTGGTGAATATCGGTTTAAAGAGTCAAGTATTTC 766
DB 2293 GTATATGAGGCCCCCAGGTGCAGCTGGTGAATTAATCTTCAAGGGCATCAAGTACTTG 2352
QY 767 CGCTTGAATACCTAGCTCTCTAGGTTATGTGGTTGTAGTGATAGACAACAGGGGATCC 826
DB 2353 CGGCTCAACACACTGGGCTCTCCCTGGGCTACGCCGTGGTGTGATTCACGCGCAGGGCTCC 2412
QY 827 TGTCAACGAGGCTTAAATTTGAAGGCGCCTTTTAAATATAAAATGGGTCAAAATAGAAATT 886
```


Db 2413 TGTGAGGAGGCTTCGGTTCGAGGGGCTTGAAACCAATGGGCCAGGTGGAGATC 2472
Qy 887 GAGCATCAGGTGGAAGACTCCAAATATCTAGTCTTCGATATGATTTCAATGACTTAGAT 946
Db 2473 GAGGACAGGTGGAGGGCTTCGAGTTCGTGGCCGAGAAATATGGCTTCATCGACCTGAGC 2532
Qy 947 CGTGTGGGCATCCACCGCTGGTCTATGAGAGATACCTTCCTCGATGGCAATTAATCGAG 1006
Db 2533 CGAGTTGCCATCCATGCTGGTCTTAAGGGGGCTTCCTCTGCTCATGGGCTTAATCCAC 2592
Qy 1007 AGGTGAGATATCTTCAGGGTTCCTATTTGCTGGGGCCCGAGTCACTCTCTGTGGATCTTAT 1066
Db 2593 AAGCCCGAGGTTCAGGTGGCCATCGCGGGTGCCCCGGTCAACGCTCTGGATGGCCTAC 2652
Qy 1067 GATACAGGATACACGGA 1083
Db 2653 GACACAGGTACACTGA 2669
RESULT 14
US-09-976-674-24
; Sequence 24, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976, 674
; CURRENT FILING DATE: 2001-10-12
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 24
; LENGTH: 4302
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-976-674-24
Query Match 31.8%; Score 344.2; DB 3; Length 4302;
Best Local Similarity 64.5%; Pred. No. 3.8e-91;
Matches 514; Conservative 0; Mismatches 283; Indels 0; Gaps 0;
Qy 287 AGGATCATAGTTGATGAAGTCAGAGGCTGGTATATTTGAAGGCCACCAAGACTCCCT 346
Db 1873 AAGATCTGGGTCAATGAGGAGACCAAGCTGGTGTACTTCCAGGGCACCAGGACACGCG 1932
Qy 347 TTAGAGCATCACCTGATAGTACGTAAATCTCGGAGAGGTGACAAGGTGACT 406
Db 1933 CTGGAGCACCACTTACGTGGTACGTATGAGGCGCGCGGAGATCGTAGCCTAC 1992
Qy 407 GAACGTGCTACTCACATCTTCTGTCATCAGTCAGTCACTGCTCTTTATTAAGTAAG 466
Db 1993 ACGCCGGCTTCTCCATAGTCTGCTCATGAGCCAGAACTTCGACATGTTGTCAGCCAC 2052
Qy 467 TAATAGTAACAGAGAATCCACACTGTGTGTCCTTTTCAAGCTATCAAGTCTCGAAGAT 526
Db 2053 TACAGCAGCGTGAGCAGCGCGCTCGGTGTCAGCTTACAACTGAGCGGGCCCGACGAC 2112
Qy 527 GACCCCACTTCCAAACAAGGAATTTTGGGCCACCATTTTGGATTTCAGCAGGTCTCTT 586
Db 2113 GACCCCTGCAACGAGCGCCCGCTTCTGGGCTAGCATGATGAGGCGAGCCAGTGCCTT 2172
Qy 587 CTGACTACTACTCTCCAGAAATTTCTCTTTTGAAGTACTACTGGATTTTACATGTAT 646
Db 2173 CGGATATATGTTCTCCAGAGATCTTCCATTTCCACACGCGTCTGGATGCGGCTCTAC 2232
Qy 647 GGGATGCTCTCAAGCCTCATGATCTATACGCTTGGAAAGAAATATCTTACTGTGCTGTT 706

Db 2233 GGCAATGATCTACAAGCCGCCACGCTTTCAGCCAGGGAAGACACCCACCGTCTCTTT 2292
Qy 707 ATATATGCTGCTCCTCAGGTGCAAGTGGTGAATAATCGTTTAAAGGAGTCAAGTATTTTC 766
Db 2293 GTATATGAGGCCCCCAGGTGCAAGTGGTGAATAATCTCTTCAAAGGCATCAAGTACTTG 2352
Qy 767 CGCTTGAATACCTTAGCCTCTCTAGTTATGTTGTTAGTATAGACAAACAGGGGATCC 826
Db 2353 CGGCTCAACACACTTGGCTCCCTGGCTACGCGTGGTGTGATTTGACGGCAGGGGCTCC 2412
Qy 827 TGTCAACGAGGCTTAAATTTGAAGGCGCTTTAAATATAAATGCTCAATAGAAAT 886
Db 2413 TGTCAAGGAGGCTTCGTTTCGAAGGGGCTTGAACCAATGGGCCAGTGGAGATC 2472
Qy 887 GACGATCAGGTGGAAGGACTCCAAATATCTAGTCTTCGATATGATTTTCATTTGACTTAGAT 946
Db 2473 GAGGACAGGTGGAGGCTGCGATTCGTGGCCGAGAGTATGGCTTCATCGACCTGAGC 2532
Qy 947 CGTGTGGGCATCCACCGCTGGTCTTATGAGAGATACCTCTCCCTGATGGCATTATGAGC 1006
Db 2533 CGAGTTGCCATCCATGGCTGCTTACGGGGGCTTCTCTCGCTCATGGGGCTAATCCAC 2592
Qy 1007 AGGTGAGATATCTTCAGGGTTCCTATTTGCTGGGGCCCGAGTCACTCTGTGGATCTTCTAT 1066
Db 2593 AAGCCCGAGGTTCAGGTGGCCATCGCGGGTGCCCCGGTCAACGCTCTGGATGGCCTAC 2652
Qy 1067 GATACAGGATACACGGA 1083
Db 2653 GACACAGGTACACTGA 2669
RESULT 15
US-09-976-674-36
; Sequence 36, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976, 674
; CURRENT FILING DATE: 2001-10-12
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 36
; LENGTH: 4180
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-976-674-36
Query Match 30.5%; Score 329.8; DB 3; Length 4180;
Best Local Similarity 64.7%; Pred. No. 6.7e-87;
Matches 490; Conservative 0; Mismatches 267; Indels 0; Gaps 0;
Qy 327 AAGGCCACCAAGACTCCCTTTAGAGCATCACCTGTACGTAGTACGTAAATTCCTG 386
Db 1874 AGGGACCAACAGACACGCGCTGGAGCACACCTCTACGTGGTACGTATGAGCGGCG 1933
Qy 387 GAGAGTGACAAGGCTGACTGACCGTGGGTACTCAATTTCTGCTGCTCAGTCACT 446
Db 1934 GCGAGATCGTACGCTCACACGCGCTTCTCCATAGTCTCTCCATGAGCAGAACT 1993
Qy 447 GTGACTTCTTTATAGTAAGTATAGTAAACAGAGATCCACACTGTGTGCTCTTACA 506
Db 1994 TCGACATGTTCTGAGGCGCTTACAGAGCGTGGAGCGCGCTCGGCTCGGCTGCTACA 2053
Qy 507 AGCTATCAAGTCTCTGAAGATGACCCAACTTTCGCAAAACAAAGGAATTTTGGGCCACCATTT 566
Db 2054 AGCTGAGGCGCGCGACGACGACGCCCTCTGCACAGGACGCCCGCTTCTGGGCTAGCATGA 2113

QY 567 TGGATTACAGAGGCTCTTCTTCTGACTATATCTCTCCAGAAATTTTCTTTTGAAGTA 626
Db |||||
QY 2114 TGGAGGCAGCCAGCTGCCCGGATTATGTTCTCCAGAGATCTTCCATTTCCACACGC 2173
Db |||||
QY 627 CTACTGGATTTTACATTGTTATGGGATGCTCTACAGGCTCATGATCTACAGCCTGGAAGA 686
Db |||||
QY 2174 GCTCGGATGTCCGCTCTACGGCATGATCTACAGGCCCCACGCCCTTGAGCCAGGGAAGA 2233
Db |||||
QY 687 AATATCCTTACTGTCTGTTCTATATATGTTGGTCTCTCAGGTGAGTTGGTGAATATCGGT 746
Db |||||
QY 2234 AGCACCCACCGTCTCTTTGTATATGGAGGCCCCCGGTGAGTTGATTAATCTCCT 2293
Db |||||
QY 747 TTAAGGAGTCAAGTATTTCCGCTTGAATACCTAGCCTCTCTAGGTTATGTTGTGTAG 806
Db |||||
QY 2294 TCAAGGCATCAAGTACTTTCGGCTCAACACACTGGCCTCCCTGGGCTACGCCGTTGGTTG 2353
Db |||||
QY 807 TGATAGACAACAGGGGATCTGTCTACCGAGGGCTTAAATTTGAAGGCGCCTTTAAATATA 866
Db |||||
QY 2354 TGATTGAGCGGAGGGGCTCTGTCTCAGCGAGGGCTTCGGTTGAGAGGGGCCCTGAAAAACC 2413
Db |||||
QY 867 AATGGGTCAAAATAGAAATTGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGAT 926
Db |||||
QY 2414 AATGGGCCAGGTGGAGATCGAGGACCGGTGGAGGGCTTCGAGTTCTGTCGCCGAGAAGT 2473
Db |||||
QY 927 ATGATTTCAITGACTTAGATTCGTGTGGGCATCCAGGCTGGTCTTATGGAGGATACCTCT 986
Db |||||
QY 2474 ATGGCTTTCATCGACCTGAGCGGAGTTGCCATCCATGGGTGGTCTTACGGGGCTTCTCT 2533
Db |||||
QY 987 CCCTGATGGCATTAAATGACAGGTCAGATATCTTCAGGGTTGCTATTGCTGGGGCCCCAG 1046
Db |||||
QY 2534 CGCTCATGGGGCTAAATCCACAAGCCCCAGGTGTTCAAGGTGGCCATCGGGGTGCCCGG 2593
Db |||||
QY 1047 TCACCTCTGTGGATCTTCTATGATACAGGATACACGGA 1083
Db |||||
QY 2594 TCACCGTCTGGATGGCTACGACACAGGGGTACACTGA 2630
Db |||||

Search completed: May 3, 2006, 02:13:00
Job time : 347.142 secs

GenCore version 5.1.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 4, 2006, 03:41:02 ; Search time 963.428 Seconds
(without alignments)
9295.694 Million cell updates/sec

Title:
Perfect score: 1083
Sequence: 1 ggaagagatgccagatcag.....tatgatacaggatacacgga 1083

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA_Main:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Length	DB	ID	Description
1	1083	100.0	1083	8	US-10-825-632-8	Sequence 8, Appli
2	1055.8	97.5	4309	3	US-09-976-674-14	Sequence 14, Appl
3	1055.8	97.5	4309	3	US-10-982-512-14	Sequence 14, Appl
4	791.2	73.1	2649	6	US-10-054-776-1	Sequence 1, Appli
5	791.2	73.1	2649	6	US-10-170-789-39	Sequence 39, Appl
6	791.2	73.1	3106	7	US-10-311-035-30	Sequence 30, Appl
7	791.2	73.1	3120	7	US-10-415-122-5	Sequence 5, Appli
8	791.2	73.1	3120	7	US-10-825-632-2	Sequence 2, Appli
9	791.2	73.1	3143	6	US-10-170-789-37	Sequence 37, Appl
10	789.6	72.9	2671	3	US-09-976-674-2	Sequence 2, Appli
11	789.6	72.9	2671	3	US-10-982-512-2	Sequence 12, Appl
12	779.2	71.9	4829	3	US-09-976-674-12	Sequence 12, Appl
13	779.2	71.9	4829	3	US-10-982-512-12	Sequence 12, Appl
14	581.8	53.7	2510	7	US-10-275-505-16	Sequence 16, Appl
15	581.8	53.7	2510	10	US-11-140-224-16	Sequence 16, Appl
16	497.2	45.9	4685	3	US-09-976-674-22	Sequence 22, Appl
17	497.2	45.9	4685	3	US-10-982-512-22	Sequence 22, Appl
18	479.2	44.2	4676	3	US-09-976-674-20	Sequence 20, Appl
19	479.2	44.2	4676	9	US-10-982-512-20	Sequence 20, Appl
20	459	42.4	1197	8	US-10-825-632-4	Sequence 4, Appli
21	434.2	40.1	1669	8	US-10-825-632-6	Sequence 6, Appli
22	434.2	40.1	2830	9	US-10-956-157-2177	Sequence 2177, Ap
23	428.2	39.5	4523	3	US-09-976-674-8	Sequence 8, Appli

RESULT 1

US-10-825-632-8
; Sequence 8, Application US/10825632
; Publication No. US20040191826A1
; GENERAL INFORMATION:
; APPLICANT: ABBOTT, Catherine Anne
; APPLICANT: GORRELL, Mark Douglas
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES
; FILE REFERENCE: FCSB-100-Div. 1
; CURRENT APPLICATION NUMBER: US/10/825,632
; CURRENT FILING DATE: 2004-04-15
; PRIOR APPLICATION NUMBER: US 10/070,464
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: PCT/AU00/01085
; PRIOR FILING DATE: 2000-09-11
; PRIOR APPLICATION NUMBER: AU PQ5709
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: AU PQ2762
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 1083
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-10-825-632-8

Query Match	100.0%;	Score 1083;	DB 8;	Length 1083;
Best Local Similarity	100.0%;	Pred. No. 2.2e-304;	Mismatches 0;	Indels 0;
Matches 1083;	Conservative 0;			Gaps 0;
QY	1	GGAGAAGATCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAAAGAAATTTGATAG	60	
Db	1	GGAGAAGATCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAAAGAAATTTGATAG	60	
QY	61	ATATTTCTGGCTATTGGTGGTGTCCAAAAGCTGAAACAACTCCACAGTGGTGGTAAATTTCT	120	
Db	61	ATATTTCTGGCTATTGGTGGTGTCCAAAAGCTGAAACAACTCCACAGTGGTGGTAAATTTCT	120	
QY	121	TAGAATTTCTATGAGAGAAATGATGAATCTGAGTGGGAAATTTATTCATGTTACATCCCC	180	
Db	121	TAGAATTTCTATGAGAGAAATGATGAATCTGAGTGGGAAATTTATTCATGTTACATCCCC	180	
QY	181	TATGTTGGAACAAGAGGAGGGCAGATTCATTCGGTTATCTTAAACAGGTACAGCAATCC	240	
Db	181	TATGTTGGAACAAGAGGAGGGCAGATTCATTCGGTTATCTTAAACAGGTACAGCAATCC	240	

241 TAAAGTCACATTTTAAAGATGTGAGAAATATGATTGCTGGAAGGAAGATCATAGTTGA 300
Db TAAAGTCACATTTTAAAGATGTGAGAAATATGATTGCTGGAAGGAAGATCATAGTTGA 300
Qy TGAAGTCAGAGGCTGCTGATATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACCT 360
Db TGAAGTCAGAGGCTGCTGATATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACCT 360
Qy GTACGTAGTCAGTTTACGTAAATCCTGGAGAGGTGACCAAGGCTGACTGAGCCGTGCTACTC 420
Db GTACGTAGTCAGTTTACGTAAATCCTGGAGAGGTGACCAAGGCTGACTGAGCCGTGCTACTC 420
Qy ACATTTCTGTCGATCAGTCAGCACTGTGACTTCTTTTAAAGTAAGTATAGTATACCCAGAA 480
Db ACATTTCTGTCGATCAGTCAGCACTGTGACTTCTTTTAAAGTAAGTATAGTATACCCAGAA 480
Qy GAATCCACACTGTGTGCTCTTTTACAAGCTATCAAGTCTCTGAAGATGACCCAACTTGCAA 540
Db GAATCCACACTGTGTGCTCTTTTACAAGCTATCAAGTCTCTGAAGATGACCCAACTTGCAA 540
Qy AACAAAGGAATTTTGGGCCACCAATTTTGGATTACAGAGTCTCTTCTTCTGACTATACCTC 600
Db AACAAAGGAATTTTGGGCCACCAATTTTGGATTACAGAGTCTCTTCTTCTGACTATACCTC 600
Qy TCCAGAAATTTTCTCTTTTGAAGTACTACTGGAATTTACATTTGATGGGATGCTCTACAA 660
Db TCCAGAAATTTTCTCTTTTGAAGTACTACTGGAATTTACATTTGATGGGATGCTCTACAA 660
Qy GCCTCATGATCTACAGCCTTGGAAAGAAATATCTACTGTGCTGTTTATATGATGGTGGTCC 720
Db GCCTCATGATCTACAGCCTTGGAAAGAAATATCTACTGTGCTGTTTATATGATGGTGGTCC 720
Qy TCAGGTGAGTGGTGAATTAATCGGTTTAAAGAGTCAAGTATTTCCGCTTTGAATACCTT 780
Db TCAGGTGAGTGGTGAATTAATCGGTTTAAAGAGTCAAGTATTTCCGCTTTGAATACCTT 780
Qy AGCCTCTAGTGTATGTTGCTGTAGTATAGACACAGGGGATCCTCTCACCGAGGCT 840
Db AGCCTCTAGTGTATGTTGCTGTAGTATAGACACAGGGGATCCTCTCACCGAGGCT 840
Qy TAAATTTGAAGGCGCTTTAAATATAAATGGTCAATAGAAATGAGATCAGGTGGA 900
Db TAAATTTGAAGGCGCTTTAAATATAAATGGTCAATAGAAATGAGATCAGGTGGA 900
Qy AGGACTCCAATATCTAGCTTCTCGATATGATTTCAATGACCTTAGATCCTGTGGGCATCCA 960
Db AGGACTCCAATATCTAGCTTCTCGATATGATTTCAATGACCTTAGATCCTGTGGGCATCCA 960
Qy CGGCTGCTCTATGGAGGATACCTCTCCCTGATGGCATTTAATGCAAGGTGAGATATCTT 1020
Db CGGCTGCTCTATGGAGGATACCTCTCCCTGATGGCATTTAATGCAAGGTGAGATATCTT 1020
Qy CAGGTTGCTATTGCTGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACAC 1080
Db CAGGTTGCTATTGCTGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACAC 1080
Qy 1081 GGA 1083
Db 1081 GGA 1083

RESULT 2
US-09-976-674-14
; Sequence 14, Application US/09976674
; Patent No. US2020115943A1
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669

; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
; LENGTH: 4309
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-976-674-14
Query Match 97.5%; Score 1055.8; DB 3; Length 4309;
Best Local Similarity 99.5%; Pred. No. 4.2e-296;
Matches 1080; Conservative 0; Mismatches 2; Indels 3; Gaps 2;
Qy 1 GGAAGAAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCCAAGAAAGATTTGATAG 60
Db 942 GGAAGAAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCCAAGAAAGATTTGATAG 1001
Qy 61 ATATTCTGGCTATTGGTGGTGTCCAAAGCTGAAACAACTCCCAAGTGGTGGTAAATTTCT 120
Db 1002 ATATTCTGGCTATTGGTGGTGTCCAAAGCTGAAACAACTCCCAAGTGGTGGTAAATTTCT 1061
Qy 121 TAGAATTTCTATATGAAGAAATGATGAATCTGAGGTGGAAATTTATTCATGTTTACATCC 180
Db 1062 TAGAATTTCTATATGAAGAAATGATGAATCTGAGGTGGAAATTTATTCATGTTTACATCC 1121
Qy 181 TATGTTGGAACAAGAGGAGGCAGATTCATTCGGTATCTCTAAACAGAGTACAGCAATCC 240
Db 1122 TATGTTGGAACAAGAGGAGGCAGATTCATTCGGTATCTCTAAACAGAGTACAGCAATCC 1181
Qy 241 TAAAGTCACATTTTAAAGATGTGAGAAATATGATTGATGCTGAAGGAAGATCATAGTTGA 300
Db 1182 TAAAGTCACATTTTAAAGATGTGAGAAATATGATTGATGCTGAAGGAAGATCATAGTTGA 1240
Qy 301 TGAAGTCAGAGGCTGCTGATATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACCT 360
Db 1241 TGAAGTCAGAGGCTGCTGATATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACCT 1300
Qy 361 GTACGTAGTCAGTTTACGTAAATCCTGGAGAGGTGACCAAGGCTGACTGACCGTGGCTACTC 420
Db 1301 GTACGTAGTCAGTTTACGTAAATCCTGGAGAGGTGACCAAGGCTGACTGACCGTGGCTACTC 1360
Qy 421 ACATTTCTGTCGATCAGTCAGCACTGTGACTTCTTTTAAAGTAAGTATAGTATACCCAGAA 480
Db 1361 ACATTTCTGTCGATCAGTCAGCACTGTGACTTCTTTTAAAGTAAGTATAGTATACCCAGAA 1420
Qy 481 GAATCCACACTGTGTGCTCTTTTACAAGCTATCAAGTCTCTGAAGATGACCCAACTTGCAA 540
Db 1421 GAATCCACACTGTGTGCTCTTTTACAAGCTATCAAGTCTCTGAAGATGACCCAACTTGCAA 1480
Qy 541 AACAAAGGAATTTTGGGCCACCAATTTTGGATTACAGAGTCTCTTCTGACTATACCTC 600
Db 1481 AACAAAGGAATTTTGGGCCACCAATTTTGGATTACAGAGTCTCTTCTGACTATACCTC 1540
Qy 601 TCCAGAAATTTTCTCTTTTGAAGTACTACTGGAATTTACATTTGATGGGATGCTCTACAA 660
Db 1541 TCCAGAAATTTTCTCTTTTGAAGTACTACTGGAATTTACATTTGATGGGATGCTCTACAA 1600
Qy 661 GCCTCATGATCTACAGCCTTGGAAAGAAATATCTACTGTGCTGTTTATATGATGGTGGTCC 718
Db 1601 GCCTCATGATCTACAGCCTTGGAAAGAAATATCTACTGTGCTGTTTATATGATGGTGGTCC 1660
Qy 719 CCTCAGGTGAGTGGTGAATTAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTTGAATACC 778
Db 1661 CCTCAGGTGAGTGGTGAATTAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTTGAATACC 1720
Qy 779 CTAGCCTCTCTAGGTTATGTTGTTAGTATAGACCAACAGGGGATCCTGTCAACCGAGGG 838
Db 1721 CTAGCCTCTCTAGGTTATGTTGTTAGTATAGACCAACAGGGGATCCTGTCAACCGAGGG 1780
Qy 839 CTTAAATTTGAAGGCGCTTTAAATATAAATGGGTCAATAGAAATTCACGATCAGGTG 898

Db 1781 CTTAAATTTGAAGCGCCCTTAAATATAAAATGGGTCAAAATAGAAATTTGACGATCAGGTG 1840
Qy 899 GAAGGACTCAATATCTAGCTTCTCGATATGATTTTCATTGACTTAGATCGTGTGGGCATC 958
Db 1841 GAAGGACTCAATATCTAGCTTCTCGATATGATTTTCATTGACTTAGATCGTGTGGGCATC 1900
Qy 959 CAGGCTGGTCTATCGAGGATACCTTCCCTGATGGCATTAATGCGAGGTGATGATC 1018
Db 1901 CAGGCTGGTCTATCGAGGATACCTTCCCTGATGGCATTAATGCGAGGTGATGATC 1960
Qy 1019 TTCAAGGTTGCTATTTGCTGGGGCCCGAGTCACTCTGTGATCTTCTATGATACAGGATAC 1078
Db 1961 TTCAAGGTTGCTATTTGCTGGGGCCCGAGTCACTCTGTGATCTTCTATGATACAGGATAC 2020
Qy 1079 ACGGA 1083
Db 2021 ACGGA 2025

RESULT 3
US-10-982-512-14
; Sequence 14, Application US/10982512
; Publication No. US20050059081A1
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akineanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/10/982,512
; CURRENT FILING DATE: 2004-11-05
; PRIOR APPLICATION NUMBER: US/09/976,674
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
; LENGTH: 4309
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-982-512-14

Query Match 97.5%; Score 1055.8; DB 9; Length 4309;
Best Local Similarity 99.5%; Pred. No. 4.2e-296;
Matches 1080; Conservative 0; Mismatches 2; Indels 3; Gaps 2;
Qy 1 GGAAGAAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAAGAGAAATTTGATAG 60
Db 942 GGAAGAAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAAGAGAAATTTGATAG 1001
Qy 61 ATATTCGGCTATTTGGTGTGTCCTCAAAAGCTGAACAACCTCCAGTGGTGTGTAATTTCT 120
Db 1002 ATATTCGGCTATTTGGTGTGTCCTCAAAAGCTGAACAACCTCCAGTGGTGTGTAATTTCT 1061
Qy 121 TAGAATTTCTATATGAGAAATGATGAATCTGAGGTGGAATTTATTCATGTTACATCCCC 180
Db 1062 TAGAATTTCTATATGAGAAATGATGAATCTGAGGTGGAATTTATTCATGTTACATCCCC 1121
Qy 181 TATGTTGGAACCAAGAGGCGAGATTCATTCGGTTATCTCTAAACAGGTACAGCAAAATCC 240
Db 1122 TATGTTGGAACCAAGAGGCGAGATTCATTCGGTTATCTTAAACAGGTACAGCAAAATCC 1181
Qy 241 TAAAGTCACCTTTTAAAGTGTGAGAAATATGATGCTGAAGGAGGATCATAGTTGA 300
Db 1182 TAAAGTCACCTTTTAAAGTGTGAGAAATATGATGCTGAAGGAGGATCATAGTTGA 1240
Qy 301 TGAAGTCAGAGGCTGTATATTTTGAAGGACCAAGACTCCCTTTAGAGCATCACCT 360
Db 1241 TGAAGTCAGAGGCTGTATATTTTGAAGGACCAAGACTCCCTTTAGAGCATCACCT 1300

Qy 361 GTACGTAGTCAGTTACGTAAATCTCGAGAGAGGTGA CAAGGCTGACTGACCGTGGCTACTC 420
Db 1301 GTACGTAGTCAGTTACGTAAATCTCGAGAGAGGTGA CAAGGCTGACTGACCGTGGCTACTC 1360
Qy 421 ACATTCCTGTCATCAGTCAGCAGCTGTCTCTCTTTTATAAGTAAGTATAGTAACAGAA 480
Db 1361 ACATTCCTGTCATCAGTCAGCAGCTGTCTCTCTTTTATAAGTAAGTATAGTAACAGAA 1420
Qy 481 GAATCCACACTGTGTGTGCCCTTTTACAAAGGTATCAAGTCTGAAGATGACCCAACTTTGCAA 540
Db 1421 GAATCCACACTGTGTGTGCCCTTTTACAAAGGTATCAAGTCTGAAGATGACCCAACTTTGCAA 1480
Qy 541 AACAAAGGAATTTTGGGCCACCAATTTTGGATTGAGAGGTCTCTCTCTGACTATACCTCC 600
Db 1481 AACAAAGGAATTTTGGGCCACCAATTTTGGATTGAGAGGTCTCTCTCTGACTATACCTCC 1540
Qy 601 TCCAGAAATTTTCTCTTTTGAAGTACTACTGGAATTTACATTTGATGGAATGCTCTACAA 660
Db 1541 TCCAGAAATTTTCTCTTTTGAAGTACTACTGGAATTTACATTTGATGGAATGCTCTACAA 1600
Qy 661 GCCTCATGATCTACAGCCTGGAAAGAAATATCTCTACTGTGCTGTTCATATATGTTGGTGG--T 718
Db 1601 GCCTCATGATCTACAGCCTGGAAAGAAATATCTCTACTGTGCTGTTCATATATGTTGGTGGTCT 1660
Qy 719 CCTCAGGTGCAAGTTGGTGAATATCGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACC 778
Db 1661 CCTCAGGTGCAAGTTGGTGAATATCGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACC 1720
Qy 779 CTAGCCTCTCTAGGTTATGTTGGTTGAGTGATAGACAACAGGGGATCTGTCCACGAGGG 838
Db 1721 CTAGCCTCTCTAGGTTATGTTGGTTGAGTGATAGACAACAGGGGATCTGTCCACGAGGG 1780
Qy 839 CTTAAATTTGAAGCGCCCTTTAAATATAAAATGAGTCAAAATAGAAATTCAGCATCAGGTG 898
Db 1781 CTTAAATTTGAAGCGCCCTTTAAATATAAAATGAGTCAAAATAGAAATTCAGCATCAGGTG 1840
Qy 899 GAAGGACTCCAATATCTAGCTTCTCGATATGATTTTCATTGACTTAGATCGTGTGGGCATC 958
Db 1841 GAAGGACTCCAATATCTAGCTTCTCGATATGATTTTCATTGACTTAGATCGTGTGGGCATC 1900
Qy 959 CAGGCTGGTCTTATGAGGATACCTCTCCCTGATGGGCAATTAATGAGAGGTGATATC 1018
Db 1901 CAGGCTGGTCTTATGAGGATACCTCTCCCTGATGGGCAATTAATGAGAGGTGATATC 1960
Qy 1019 TTCAGGCTGCTATTTGCTGGGGCCCGAGTCACTCTGTGATCTCTTATGATACAGGATAC 1078
Db 1961 TTCAGGCTGCTATTTGCTGGGGCCCGAGTCACTCTGTGATCTCTTATGATACAGGATAC 2020
Qy 1079 ACGGA 1083
Db 2021 ACGGA 2025

RESULT 4
US-10-054-776-1
; Sequence 1, Application US/10054776
; Publication No. US20030165818A1
; GENERAL INFORMATION:
; APPLICANT: Mark Robert Edbrooke
; APPLICANT: Alan Peter Lewis
; TITLE OF INVENTION: NOVEL PROTEIN
; FILE REFERENCE: OGI042US
; CURRENT APPLICATION NUMBER: US/10/054,776
; CURRENT FILING DATE: 2002-01-23
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 2649
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2649)

QY 340 CTCCTCTTTAGAGCATCACCTGTACGTAGTACGTAAATCCTGGAGGTGACAAG 399
Db 1587 CTCCTCTTTAGAGCATCACCTGTACGTAGTACGTAAATCCTGGAGGTGACAAG 1646
QY 400 GCTGACTGACCGTGGCTACTACATTTCTTGTGCTGATCAGTACGACCTGCTCTTTAT 459
Db 1647 GCTGACTGACCGTGGCTACTACATTTCTTGTGCTGATCAGTACGACCTGCTCTTTAT 1706
QY 460 AAGTAAGTATAGTAACCAAGAAATCCACATGCTGTCCTTTTACAAGCTATCAAGTCC 519
Db 1707 AAGTAAGTATAGTAACCAAGAAATCCACATGCTGTCCTTTTACAAGCTATCAAGTCC 1766
QY 520 TGAAGATGACCCAACTTGCACAAACAAAGAAATTTTGGGCCCAATTTTGGATTCAGCAGG 579
Db 1767 TGAAGATGACCCAACTTGCACAAACAAAGAAATTTTGGGCCCAATTTTGGATTCAGCAGG 1826
QY 580 TCCTCTTCTGACTATCTCTCCAGAAATTTTCTTTTGAAGTACTACTGATTTAC 639
Db 1827 TCCTCTTCTGACTATCTCTCCAGAAATTTTCTTTTGAAGTACTACTGATTTAC 1886
QY 640 ATTCTATGGGATGCTCTACAAGCTCATGATCTACAGCTGGAAAGAAATATCTCTACTGT 699
Db 1887 ATTCTATGGGATGCTCTACAAGCTCATGATCTACAGCTGGAAAGAAATATCTCTACTGT 1946
QY 700 GCTCTTCATATATGGTGGTCTCTCAGGTGCGATTTGGTGAATATCAAGAGTCAA 759
Db 1947 GCTCTTCATATATGGTGGTCTCTCAGGTGCGATTTGGTGAATATCAAGAGTCAA 2006
QY 760 GTATTTCCGTTGAATACCTAGCCTCTCTAGGTTATGGTGTAGTGATAGACAAG 819
Db 2007 GTATTTCCGTTGAATACCTAGCCTCTCTAGGTTATGGTGTAGTGATAGACAAG 2066
QY 820 GGGATCCTGTACCGGGGCTTAAATTTGAAGCGCTTTAAATATAAAATGGGTCAAAT 879
Db 2067 GGGATCCTGTACCGGGGCTTAAATTTGAAGCGCTTTAAATATAAAATGGGTCAAAT 2126
QY 880 AGAAATGACGATCAGGTGAAGGACTCCAATATCTAGCTTCTCGATATGATTTCAITGA 939
Db 2127 AGAAATGACGATCAGGTGAAGGACTCCAATATCTAGCTTCTCGATATGATTTCAITGA 2186
QY 940 CTTAGATCGTGTGGGATCCACGGCTGGTCTATGAGGATACCTCTCCCTGATGGCAAT 999
Db 2187 CTTAGATCGTGTGGGATCCACGGCTGGTCTATGAGGATACCTCTCCCTGATGGCAAT 2246
QY 1000 AATCAGAGGTGAGATATCTTACGGGTGCTATTGCTGGGGCCCAAGTCACTCTGTGGAT 1059
Db 2247 AATCAGAGGTGAGATATCTTACGGGTGCTATTGCTGGGGCCCAAGTCACTCTGTGGAT 2306
QY 1060 CTTCTATGATACAGGATACCGGA 1083
Db 2307 CTTCTATGATACAGGATACCGGA 2330

RESULT 6

US-10-311-035-30.
; Sequence 30, Application US/10311035
; Publication No. US20040023243A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: YUE, Henry
; APPLICANT: ELLIOTT, Vicki
; APPLICANT: GANDHI, Ameena R.
; APPLICANT: LAL, Preeti
; APPLICANT: AU-YOUNG, Janice
; APPLICANT: TRIBOULEY, Catherine M.
; APPLICANT: DELEGEANE, Angelo M.
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: NGUYEN, Danniell B.
; APPLICANT: LEE, Ernestine A.
; APPLICANT: HAFALIA, April
; APPLICANT: KHAN, Farrah A.
; APPLICANT: CHAWLA, Narinder K.

; APPLICANT: YAO, Monique G.
; APPLICANT: LU, Dyung Aina M.
; APPLICANT: ARVIZU, Chandra S.
; APPLICANT: TANG, Y. Tom
; APPLICANT: WALSH, Roderick T.
; APPLICANT: AZIMZAI, Yaida
; APPLICANT: LU, Yan
; APPLICANT: RAMKUMAR, Jayalaximi
; APPLICANT: XU, Yuming
; APPLICANT: REDDY, Roopa
; APPLICANT: DAS, Depopriya
; APPLICANT: KEARNEY, Liam
; APPLICANT: KALLICK, Deborah A.
; TITLE OF INVENTION: Proteases
; FILE REFERENCE: PI-0123 PCT
; CURRENT APPLICATION NUMBER: US/10/311,035
; CURRENT FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: 60/212,336; 60/213,955; 60/215,396; 60/216,821; 60/218,946
; PRIOR FILING DATE: 2000-06-16; 2000-06-22; 2000-06-29; 2000-07-07; 2000-07-14
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PERL Program
; SEQ ID NO 30
; LENGTH: 3106
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20040023243A1 7160544CB1
US-10-311-035-30

Query Match 73.18; Score 791.2; DB 7; Length 3106;
Best Local Similarity 99.08; Pred. No. 4.5e-219;
Matches 796; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 280 TGAAGAGGATCATAGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGGCCACCAAGA 339
Db 1729 TGGATCTAATATCCAAAGTTGATGATCAGAGGCTGGTATATTTTGAAGGCCACCAAGA 1788
QY 340 CTCCTCTTTAGAGCATCACCTGTACGTGACGTTCAGTAAATCCTGGAGGTGACAAG 399
Db 1789 CTCCTCTTTAGAGCATCACCTGTACGTGACGTTCAGTAAATCCTGGAGGTGACAAG 1848
QY 400 GCTGACTGACCGTGGCTACTACATTTCTTGTGCTGATCAGTACGACTGTCTTTAT 459
Db 1849 GCTGACTGACCGTGGCTACTACATTTCTTGTGCTGATCAGTACGACTGTCTTTAT 1908
QY 460 AAGTAAGTATAGTAACCAAGAAATCCACATGCTGTCCTTTTACAAGCTATCAAGTCC 519
Db 1909 AAGTAAGTATAGTAACCAAGAAATCCACATGCTGTCCTTTTACAAGCTATCAAGTCC 1968
QY 520 TGAAGATGACCCAACTTGCACAAACAAAGAAATTTTGGGCCCAATTTTGGATTCAGCAGG 579
Db 1969 TGAAGATGACCCAACTTGCACAAACAAAGAAATTTTGGGCCCAATTTTGGATTCAGCAGG 2028
QY 580 TCCTCTTCTGACTATCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGATTAC 639
Db 2029 TCCTCTTCTGACTATCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGATTAC 2088
QY 640 ATTGTATGGGATGCTCTACAAGCTCATGATCTACAGCTGGAAAGAAATATCTCTACTGT 699
Db 2089 ATTGTATGGGATGCTCTACAAGCTCATGATCTACAGCTGGAAAGAAATATCTCTACTGT 2148
QY 700 GCTGTTTCATATATGGTGGTCTCTCAGGTGCTAGTGGTGAATATCGGTTTAAAGGAGTCAA 759
Db 2149 GCTGTTTCATATATGGTGGTCTCTCAGGTGCTAGTGGTGAATATCGGTTTAAAGGAGTCAA 2208
QY 760 GTATTTCCGCTTGAATACCTTAGCCTCTCTAGGTTATGCTGGTGTAGTGATAGACAAG 819
Db 2209 GTATTTCCGCTTGAATACCTTAGCCTCTCTAGGTTATGCTGGTGTAGTGATAGACAAG 2268
QY 820 GGGATCCTGTCAACCGAGGGCTTAAATTTGAAGCGCTTTAAATATAAAATGGGTCAAAT 879
Db 2269 GGGATCCTGTCAACCGAGGGCTTAAATTTGAAGCGCTTTAAATATAAAATGGGTCAAAT 2328

```
QY 880 AGAAATTGACGATCAGGTGGAAGACATCCCAATATCTAGCTTCTCGATATGATTTTCATTGA 939
Db 2329 AGAAATTGACGATCAGGTGGAAGACATCCCAATATCTAGCTTCTCGATATGATTTTCATTGA 2388
QY 940 CTTAGATCGTGTGGGCATCCACGGCTGGTCTCTATGAGGATACCTCTCCCTGATGGCAAT 999
Db 2389 CTTAGATCGTGTGGGCATCCACGGCTGGTCTCTATGAGGATACCTCTCCCTGATGGCAAT 2448
QY 1000 AATCAGAGGTCAGATATCTTCAAGGTTGCTATGCTGGGGCCCGAGTCACTCTGTGGAT 1059
Db 2449 AATCAGAGGTCAGATATCTTCAAGGTTGCTATGCTGGGGCCCGAGTCACTCTGTGGAT 2508
QY 1060 CTTCTATGATACAGGATACACGGA 1083
Db 2509 CTTCTATGATACAGGATACACGGA 2532

RESULT 7
US-10-415-122-5
; Sequence 5, Application US/10415122
; Publication No. US20040053369A1
; GENERAL INFORMATION:
; APPLICANT: THE UNIVERSITY OF SYDNEY
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES
; FILE REFERENCE: FP15217
; CURRENT APPLICATION NUMBER: US/10/415,122
; CURRENT FILING DATE: 2003-08-07
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 3120
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-415-122-5

Query Match 73.1%; Score 791.2; DB 7; Length 3120;
Best Local Similarity 99.0%; Pred. No. 4.5e-219;
Matches 796; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 280 TGAAGGAAGATCATAGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGGCACCAAGA 339
Db 1740 TGGATCTTAATATCCAAAGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGGCACCAAGA 1799
QY 340 CTCCTCTTTAGAGCATCACCCTGTACGTAGTTCAGTTACGTAATAATCCTGAGAGGTGACAAG 399
Db 1800 CTCCTCTTTAGAGCATCACCCTGTACGTAGTTCAGTTACGTAATAATCCTGAGAGGTGACAAG 1859
QY 400 GCTGACTGACCGTGGCTACTCACATTCCTGTCGATCAGTCAGCACTGTGACTTCTTTAT 459
Db 1860 GCTGACTGACCGTGGCTACTCACATTCCTGTCGATCAGTCAGCACTGTGACTTCTTTAT 1919
QY 460 AAGTAAGTATAGTAACCAAGAAATCCACACTGTGTGTCCTTTTACAGCTATCAAGTCC 519
Db 1920 AAGTAAGTATAGTAACCAAGAAATCCACACTGTGTGTCCTTTTACAGCTATCAAGTCC 1979
QY 520 TGAAGATGACCCAACTTGCAAAACAAAGAAATTTTGGGCCACCACTTTTGGATTTCAGCAGG 579
Db 1980 TGAAGATGACCCAACTTGCAAAACAAAGAAATTTTGGGCCACCACTTTTGGATTTCAGCAGG 2039
QY 580 TCCTCTTCTGACTACTCTCTCCAGAAATTTTCTCTTTTGAAGAGTACTACTGGAATTCAC 639
Db 2040 TCCTCTTCTGACTACTCTCTCCAGAAATTTTCTCTTTTGAAGAGTACTACTGGAATTCAC 2099
QY 640 ATTGTATGGGATGCTCTACAAGCCTCATGATCTACAGCCTTGAAAGAAATATCCTACTGT 699
Db 2100 ATTGTATGGGATGCTCTACAAGCCTCATGATCTACAGCCTTGAAAGAAATATCCTACTGT 2159
QY 700 GCTGTTTCATATATGGTGGTCTCCTCAGTGCAGTTGGTGAATAATCGGTTTAAAGAGGTCMA 759
Db 2160 GCTGTTTCATATATGGTGGTCTCCTCAGTGCAGTTGGTGAATAATCGGTTTAAAGAGGTCMA 2219
QY 760 GTATTTCCGCTTGAATACCTTAGCCTCTCTAGGTTTATGTGGTTTGTGTGATGATACACAG 819
```

```
Db 2220 GTATTTCCGCTTGAATACCTTAGCCTCTCTAGGTTTATGTGGTTTGTGTGATGATACACAG 2279
QY 820 GGGATCCTGTCCACGAGGCTTAAATTTGAAGGGCTTTTAAATATAAAATGGTCAAT 879
Db 2280 GGGATCCTGTCCACGAGGCTTAAATTTGAAGGGCTTTTAAATATAAAATGGTCAAT 2339
QY 880 AGAAATTGACGATCAGGTGGAAGACATCCCAATATCTAGCTTCTCGATATGATTTTCATTGA 939
Db 2340 AGAAATTGACGATCAGGTGGAAGACATCCCAATATCTAGCTTCTCGATATGATTTTCATTGA 2399
QY 940 CTTAGATCGTGTGGGCATCCACGGCTGGTCTCTATGAGGATACCTCTCCCTGATGGCAAT 999
Db 2400 CTTAGATCGTGTGGGCATCCACGGCTGGTCTCTATGAGGATACCTCTCCCTGATGGCAAT 2459
QY 1000 AATCAGAGGTCAGATATCTTCAAGGTTGCTATGCTGGGGCCCGAGTCACTCTGTGGAT 1059
Db 2460 AATCAGAGGTCAGATATCTTCAAGGTTGCTATGCTGGGGCCCGAGTCACTCTGTGGAT 2519
QY 1060 CTTCTATGATACAGGATACACGGA 1083
Db 2520 CTTCTATGATACAGGATACACGGA 2543

RESULT 8
US-10-825-632-2
; Sequence 2, Application US/10825632
; Publication No. US20040191826A1
; GENERAL INFORMATION:
; APPLICANT: ABBOTT, Catherine Anne
; APPLICANT: GORRELL, Mark Douglas
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES
; FILE REFERENCE: FCSB-100-Div. 1
; CURRENT APPLICATION NUMBER: US/10/825,632
; CURRENT FILING DATE: 2004-04-15
; PRIOR APPLICATION NUMBER: US 10/070,464
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: PCT/AU00/01085
; PRIOR FILING DATE: 2000-09-11
; PRIOR APPLICATION NUMBER: AU PQ5709
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: AU PQ2762
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 3120
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-10-825-632-2

Query Match 73.1%; Score 791.2; DB 8; Length 3120;
Best Local Similarity 99.0%; Pred. No. 4.5e-219;
Matches 796; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 280 TGAAGGAAGATCATAGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGGCACCAAGA 339
Db 1740 TGGATCTTAATATCCAAAGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGGCACCAAGA 1799
QY 340 CTCCTCTTTAGAGCATCACCCTGTACGTAGTTCAGTTACGTAATAATCCTGAGAGGTGACAAG 399
Db 1800 CTCCTCTTTAGAGCATCACCCTGTACGTAGTTCAGTTACGTAATAATCCTGAGAGGTGACAAG 1859
QY 400 GCTGACTGACCGTGGCTACTCACATTCCTGTCGATCAGTCAGCACTGTGACTTCTTTAT 459
Db 1860 GCTGACTGACCGTGGCTACTCACATTCCTGTCGATCAGTCAGCACTGTGACTTCTTTAT 1919
QY 460 AAGTAAGTATAGTAACCAAGAAATCCACACTGTGTGTCCTTTTACAGCTATCAAGTCC 519
Db 1920 AAGTAAGTATAGTAACCAAGAAATCCACACTGTGTGTCCTTTTACAGCTATCAAGTCC 1979
QY 520 TGAAGATGACCCAACTTGCAAAACAAAGAAATTTTGGGCCACCACTTTTGGATTTCAGCAGG 579
```


QY 640 ATTGTATGGGATGCTCTACAGCCTCATGATCTACAGCCTGAAAGAAATATCTACTGT 699
DB 2115 ATTGTATGGGATGCTCTACAGCCTCATGATCTACAGCCTGAAAGAAATATCTACTGT 2174
QY 700 CTTTTCATATATGGTGGTCTCTCAGTGCAGTTGGTGAATATCGGTTTAAAGAGTCAA 759
DB 2175 CTTTTCATATATGGTGGTCTCTCAGTGCAGTTGGTGAATATCGGTTTAAAGAGTCAA 2234
QY 760 GTATTTCCGCTTGAATACCTAGCCTCTCTAGGTTATGGTTAGTGATAGTAAAGAGTCAA 819
DB 2235 GTATTTCCGCTTGAATACCTAGCCTCTCTAGGTTATGGTTAGTGATAGTAAAGAGTCAA 2294
QY 820 GGGATCTGTGCACGAGGGCTTAAATTTGAAGCGCCTTTAAATATAAAATGGGTCAAAT 879
DB 2295 GGGATCTGTGCACGAGGGCTTAAATTTGAAGCGCCTTTAAATATAAAATGGGTCAAAT 2354
QY 880 AGAAATTGACGATCAGGTGAAGGACTCCCAATATCTAGCTTCTCGATATGATTTCAATTGA 939
DB 2355 AGAAATTGACGATCAGGTGAAGGACTCCCAATATCTAGCTTCTCGATATGATTTCAATTGA 2414
QY 940 CTTAGATCGTGTGGGCATCCACGGCTGGTCTATGAGGATACCTCTCCCTGATGGCAAT 999
DB 2415 CTTAGATCGTGTGGGCATCCACGGCTGGTCTATGAGGATACCTCTCCCTGATGGCAAT 2474
QY 1000 AATGACAGGTTCAGATATCTTTCAGGGTTGCTATTGCTGGGGCCCCAGTCACTCTGTGGAT 1059
DB 2475 AATGACAGGTTCAGATATCTTTCAGGGTTGCTATTGCTGGGGCCCCAGTCACTCTGTGGAT 2534
QY 1060 CTTCTATGATACAGGATACACGGA 1083
DB 2535 CTTCTATGATACAGGATACACGGA 2558

RESULT 10

US-09-976-674-2

; Sequence 2, Application US/09976674

; Patent No. US20020115843A1

; GENERAL INFORMATION:

; APPLICANT: Qi, Steve

; APPLICANT: Akinsanya, Karen

; APPLICANT: Riviere, Pierre

; APPLICANT: Junien, Jean-Louis

; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV

; FILE REFERENCE: 70669

; CURRENT APPLICATION NUMBER: US/09/976,674

; CURRENT FILING DATE: 2001-10-12

; PRIOR APPLICATION NUMBER: US 60/240,117

; PRIOR FILING DATE: 2000-10-12

; NUMBER OF SEQ ID NOS: 61

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 2

; LENGTH: 2671

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-976-674-2

Query Match 72.9%; Score 789.6; DB 3; Length 2671;

Best Local Similarity 98.9%; Pred. No. 1.2e-218;

Matches 795; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 280 TGAAGGAAGATCATAGTTGATGAAGTCAAGAGGCTGGTATATTTTGAAGGCCACCAAGA 339
DB 1534 TGGATCTAATATCAAGTTGATGAAGTCAAGAGGCTGGTATATTTTGAAGGCCACCAAGA 1593
QY 340 CTCCTTTAGACATCACTGTACGTAGTCACTTAATCTCGAGAGTGAAGAAG 399
DB 1594 CTCCTTTAGACATCACTGTACGTAGTCACTTAATCTCGAGAGTGAAGAAG 1653
QY 400 GCTGACTGACCGTGGTACTACATCTTCTGTCATGATGAGTCACTGCTCTTTAT 459
DB 1654 GCTGACTGACCGTGGTACTACATCTTCTGTCATGATGAGTCACTGCTCTTTAT 1713
QY 460 AAGTAAGTATAGTAACACAGAGAAATCCACACTGTGTGCTCCCTTTTACAGCTATCAAGTCC 519

DB 1714 AAGTAAGTATAGTAACACAGAGAAATCCACACTGTGTGCTCCCTTTTACAGCTATCAAGTCC 1773
QY 520 TGAAGATGACCCAACTTGGCAAAAACAAAGGAATTTTGGGCCACCATTTTGGATTTCAAGCAGG 579
DB 1774 TGAAGATGACCCAACTTGGCAAAAACAAAGGAATTTTGGGCCACCATTTTGGATTTCAAGCAGG 1833
QY 580 TCTCTTCTCTGACTATATCTCTCCAGAAAATTTTCTCTTTTGAAGAGTACTACTGGATTTTAC 639
DB 1834 TCTCTTCTCTGACTATATCTCTCCAGAAAATTTTCTCTTTTGAAGAGTACTACTGGATTTTAC 1893
QY 640 ATTGTATGGGATGCTCTTACAAAGCCTCATGATCTACAGCCTGAAAGAAAATATCTCTACTGT 699
DB 1894 ATTGTATGGGATGCTCTTACAAAGCCTCATGATCTACAGCCTGAAAGAAAATATCTCTACTGT 1953
QY 700 GCTGTTTCATATATATGGTGGTCTCTCAGGTGCAAGTTGGTGAATATCGGTTTAAAGAGGATCAA 759
DB 1954 GCTGTTTCATATATATGGTGGTCTCTCAGGTGCAAGTTGGTGAATATCGGTTTAAAGAGGATCAA 2013
QY 760 GTATTTCCGCTTGAATACCTAGCCTCTCTAGGTTATGGTTAGTGATAGTAAAGAGTCAA 819
DB 2014 GTATTTCCGCTTGAATACCTAGCCTCTCTAGGTTATGGTTAGTGATAGTAAAGAGTCAA 2073
QY 820 GGGATCTGTGCACGAGGGCTTAAATTTGAAGCGCCTTTAAATATAAAATGGGTCAAAT 879
DB 2074 GGGATCTGTGCACGAGGGCTTAAATTTGAAGCGCCTTTAAATATAAAATGGGTCAAAT 2133
QY 880 AGAAATTGACGATCAGGTGAAGGACTCCCAATATCTAGCTTCTCGATATGATTTCAATTGA 939
DB 2134 AGAAATTGACGATCAGGTGAAGGACTCCCAATATCTAGCTTCTCGATATGATTTCAATTGA 2193
QY 940 CTTAGATCGTGTGGGCATCCACGGCTGGTCTATGAGGATACCTCTCCCTGATGGCAAT 999
DB 2194 CTTAGATCGTGTGGGCATCCACGGCTGGTCTATGAGGATACCTCTCCCTGATGGCAAT 2253
QY 1000 AATGACAGGTTCAGATATCTTTCAGGGTTGCTATTGCTGGGGCCCCAGTCACTCTGTGGAT 1059
DB 2254 AATGACAGGTTCAGATATCTTTCAGGGTTGCTATTGCTGGGGCCCCAGTCACTCTGTGGAT 2313
QY 1060 CTTCTATGATACAGGATACACGGA 1083
DB 2314 CTTCTATGATACAGGATACACGGA 2337

RESULT 11

US-10-982-512-2

; Sequence 2, Application US/10982512

; Publication No. US20050059081A1

; GENERAL INFORMATION:

; APPLICANT: Qi, Steve

; APPLICANT: Akinsanya, Karen

; APPLICANT: Riviere, Pierre

; APPLICANT: Junien, Jean-Louis

; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV

; FILE REFERENCE: 70669

; CURRENT APPLICATION NUMBER: US/10/982,512

; CURRENT FILING DATE: 2004-11-05

; PRIOR APPLICATION NUMBER: US/09/976,674

; PRIOR FILING DATE: 2001-10-12

; PRIOR APPLICATION NUMBER: US 60/240,117

; PRIOR FILING DATE: 2000-10-12

; NUMBER OF SEQ ID NOS: 61

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 2

; LENGTH: 2671

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-982-512-2

Query Match 72.9%; Score 789.6; DB 9; Length 2671;

Best Local Similarity 98.9%; Pred. No. 1.2e-218;

Matches 795; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy	280	TGAAGGAAGGATCATAGTTGATGAAGTCAGAAGCGCTGGTATATATTTTGAAGCCACAAGA	339
Db	1534	TGGATCTAATATCCAAAGTTGATGAAGTCAGAAGCGCTGGTATATATTTTGAAGCCACAAGA	1593
Qy	340	CTCCCTCTTTAGAGCATCAACCTGTACGTAGTCAGTTACGTAAATCTCGAGAGGTGACAAG	399
Db	1594	CTCCCTCTTTAGAGCATCAACCTGTACGTAGTCAGTTACGTAAATCTCGAGAGGTGACAAG	1653
Qy	400	GCTGACTGACCGTGGCTACTCACATTTCTTGCTCATCAGTCAAGCACTGTGACTCTCTTTAT	459
Db	1654	GCTGACTGACCGTGGCTACTCACATTTCTTGCTCATCAGTCAAGCACTGTGACTCTCTTTAT	1713
Qy	460	AAGTAAGTATAGTAACCAAGAGAATCCACACTGTGTGTCCTTTACAGACTATCAAGTCC	519
Db	1714	AAGTAAGTATAGTAACCAAGAGAATCCACACTGTGTGTCCTTTACAGACTATCAAGTCC	1773
Qy	520	TGAAGATGACCCAACTTGCAAAAACAAAGGAATTTGGGGCCACCATTTTGGATTCAGCAGG	579
Db	1774	TGAAGATGACCCAACTTGCAAAAACAAAGGAATTTGGGGCCACCATTTTGGATTCAGCAGG	1833
Qy	580	TCCTCTCTCTGACTATATCTCTCCAGAAAATTTTCTCTTTTGAAGTACTACTGTGATTTAC	639
Db	1834	TCCTCTCTCTGACTATATCTCTCCAGAAAATTTTCTCTTTTGAAGTACTACTGTGATTTAC	1893
Qy	640	ATTGTATGGGATGCTCTACAGCCTCATGATCTACAGCCTGGAAAGAAATATCTCTACTGT	699
Db	1894	ATTGTATGGGATGCTCTACAGCCTCATGATCTACAGCCTGGAAAGAAATATCTCTACTGT	1953
Qy	700	GCCTTTCATATATGGTGGTCTCTCAGTGCGAGTTGGTGAATAATTCGGTTTAAAGGAGTCAA	759
Db	1954	GCCTTTCATATATGGTGGTCTCTCAGTGCGAGTTGGTGAATAATTCGGTTTAAAGGAGTCAA	2013
Qy	760	GTATTTCCGTTGAATACCTTAGCCTCTTAGTTATATGTGTTAGTGTAGTAGATAGACAACAG	819
Db	2014	GTATTTCCGTTGAATACCTTAGCCTCTTAGTTATATGTGTTAGTGTAGTAGATAGACAACAG	2073
Qy	820	GGGATCTCTGTCAACGAGGCTTAAATTTGAAGGCGCCTTTAAATATAAAAATGGGTCAAAT	879
Db	2074	GGGATCTCTGTCAACGAGGCTTAAATTTGAAGGCGCCTTTAAATATAAAAATGGGTCAAAT	2133
Qy	880	AGAAATTGACGATCAGGTGGGAAGGACTCCAATATCTAGCTTCCGATATGATTTCAATTGA	939
Db	2134	AGAAATTGACGATCAGGTGGGAAGGACTCCAATATCTAGCTTCCGATATGATTTCAATTGA	2193
Qy	940	CTTAGATCTGTGTGGGATCCACGGCTGTCTCTATGGAGGATACCTCTCCCTGATGGCAATT	999
Db	2194	CTTAGATCTGTGTGGGATCCACGGCTGTCTCTATGGAGGATACCTCTCCCTGATGGCAATT	2253
Qy	1000	AATCGAGAGTCAAGATATCTTCAGGGTTGCTATTGCTGGGGCCCCAGTCACTCTCTGTGGAT	1059
Db	2254	AATCGAGAGTCAAGATATCTTCAGGGTTGCTATTGCTGGGGCCCCAGTCACTCTCTGTGGAT	2313
Qy	1060	CTTCTATGATACAGGATACACCGA	1083
Db	2314	CTTCTATGATACAGGATACACCGA	2337

RESULT 12

US-09-976-674-12
; Sequence 12, Application US/09976674
: Patent No. US20020115843A1

; FALCIC NO: USZ0VZ0T

: GENERAL INFORMATION:

```

: GENERAL INFORMATION:
: APPLICANT: Q1, Steve
: APPLICANT: Akinsanya, Karen
: APPLICANT: Riviere, Pierre
: APPLICANT: Junien, Jean-Louis
: TITLE OF INVENTION: NOVEL SERINE PROTEASE
: FILE REFERENCE: 706669
: CURRENT APPLICATION NUMBER: US/09/976,674
: CURRENT FILING DATE: 2001-10-12
: PRIOR APPLICATION NUMBER: US 60/240,117
: PRIOR FILING DATE: 2000-10-12
: NUMBER OF SEQ IDS: 61

```

RESULT 13

US-10-982-512-12
; Sequence 12, Application US/10982512
; Publication No. US20050059081A1
; GENERAL INFORMATION:

```
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 4829
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-976-674-12
```

Query Match 71.9%; Score 779.2; DB 3; Length 4829;

Query Match 71.5%, Score 1.52, 22.9
Best Local Similarity 98.8%: Pred. No. 1.8e-215;

BEST LOCAL SIMILARITY	98.8%	FREQ: NO: 1:98-215;	
Matches	796	Conservative	0: Mismatches
			8: Indels
			2: Gaps
			1:

QY	280	TG	AGGAGGATCATAGTGTGATGAAGTCACAAGCGCTGGTATATTTTGAAGGCACCAAGA	339
DB	1740	TGG	ATCTAATTCCAAGTTGATGAAGTCAGAAGCGCTGGTATATTTTGAAGGCACCAAGA	1799
QY	340	CT	CCCCCTTTAGAGATCACCTGTACGTAGTCAGTTACGTAAATCTCTGGAGAGGTGACAAG	399
DB	1800	CT	CCCCCTTTAGAGATCACCTGTACGTAGTCAGTTACGTAAATCTCTGGAGAGGTGACAAG	1859
QY	400	GCT	GACTGACCGGTGGCTACTACATCTCTGCTGCATCAGTCAGCACTGTGACTTCCTTAT	459
DB	1860	GCT	GACTGACCGGTGGCTACTACATCTCTGCTGCATCAGTCAGCACTGTGACTTCCTTAT	1919
QY	460	AAG	TAGTATAGTATACACAGAAGATCCACACTGTGTGTCCTTTACAAGCTATCAAGTCC	519
DB	1920	AAG	TAGTATAGTATACACAGAAGATCCACACTGTGTGTCCTTTACAAGCTATCAAGTCC	1979
QY	520	TGA	ATGATGCCAACCTTGCAAAAACAAGGAATTTGGGCCACCATTTTGGATTCAGCAGG	579
DB	1980	TGA	ATGATGCCAACCTTGCAAAAACAAGGAATTTGGGCCACCATTTTGGATTCAGCAGG	2039
QY	580	TC	CTCTTCTGACTATATCTCTCCAGAAAATTTTCTCTTTGAAAGTACTACTGGATTTAC	639
DB	2040	TC	CTCTTCTGACTATATCTCTCCAGAAAATTTTCTCTTTGAAAGTACTACTGGATTTAC	2099
QY	640	ATT	GTATGGGAGCTCTACAAGCTCATGATCTACAGCCTGGAAAAGAAATATCCTACTGT	699
DB	2100	ATT	GTATGGGAGCTCTACAAGCTCATGATCTACAGCCTGGAAAAGAAATATCCTACTGT	2159
QY	700	GCT	GTTCATATATGGTGG--TCCTCAGGTGCAGTTGGTGAATAATCGGPTTAAAGGATC	757
DB	2160	GCT	GTTCATATATGGTGGTCTCTCCAGTGCGAGTTGGTGAATAATCGGPTTAAAGGATC	2219
QY	758	AAG	TATTCGGCTTGAATACCTTAGCTCTCTAGGTATATGTGGTTGTAGTGATAGACAAC	817
DB	2220	AAG	TATTCGGCTTGAATACCTTAGCTCTCTAGGTATATGTGGTTGTAGTGATAGACAAC	2279
QY	818	AGG	GATCTGTGCACGAGGGCTTAAATTTGAAGGCGCCTTTAAATATATAAATGGGTCAA	877
DB	2280	AGG	GATCTGTGCACGAGGGCTTAAATTTGAAGGCGCCTTTAAATATATAAATGGGTCAA	2339
QY	878	AT	AGAAATTCAGCATCAGGTGGAAGGACTCAATATCTAGTCTTCGATATGATTTCAAT	937
DB	2340	AT	AGAAATTCAGCATCAGGTGGAAGGACTCAATATCTAGTCTTCGATATGATTTCAAT	2399
QY	938	GAC	TTAGATCTGTGTGGGCATCCAGGCTGGTCTTAGTGAGGATACCTCTCCCTGATGCA	997
DB	2400	GAC	TTAGATCTGTGTGGGCATCCAGGCTGGTCTTAGTGAGGATACCTCTCCCTGATGCA	2459
QY	998	TTA	TGCAGAGGTTCAGATATCTTCAGGGTGTCTATTGCTGGGGCCCCAGTCACTCTGTGG	1057
DB	2460	TTA	TGCAGAGGTTCAGATATCTTCAGGGTGTCTATTGCTGGGGCCCCAGTCACTCTGTGG	2517
QY	1058	AT	CTCTTATGATACAGGATACACGGA	1083
DB	2520	AT	CTCTTATGATACAGGATACACGGA	2545

100

RESULT 13

US-10-982-512-12
; Sequence 12, Application US/10982512
; Publication No. US20050059081A1
; GENERAL INFORMATION:

Db	1650	CCCAACTTGCACAAAGGAATTTTGGGCCACCATTTTGGATTTCAGCAGGTCCTCTTCC	1709
Qy	589	TGACTATACCTCCAGAAATTTCTCTTTTGAAGTACTACTGGATTTCACATTGTATGG	648
Db	1710	TGACTATACCTCTCCAGAAATTTCTCTTTTGAAGTACTACTGGATTTCACATTGTATGG	1769
Qy	649	GATGCTCTACAAGCCTCATGATCTACAGCCTGGAAGAAATATCTCTACTGTGCTGTTCAT	708
Db	1770	GATGCTCTACAAGCCTCATGATCTACAGCCTGGAAGAAATATCTCTACTGTGCTGTTCAT	1829
Qy	709	ATATGGTGGTCTCAGGTGCAAGTTCGTAATATCGTTTAAAGGAGTCAAGTATTTCCG	768
Db	1830	ATATGGTGGTCTCAGGTGCAAGTTCGTAATATCGTTTAAAGGAGTCAAGTATTTCCG	1889
Qy	769	CTTGAATACCTTAGCCTCTTAGGTTATGTGTTTGTAGTGATAGACAAACAGGGGATCCTG	828
Db	1890	CTTGAATACCTTAGCCTCTTAGGTTATGTGTTTGTAGTGATAGACAAACAGGGGATCCTG	1949
Qy	829	TCACCGAGGCTTAAATTTGAAGGCGCTTTAAATATATAAATGGGTCAAATAG	881
Db	1950	TCACCGAGGCTTAAATTTGAAGGCGCTTTAAATATATAAATGGGTCAAATAG	2002

RESULT 15

US-11-140-224-16

Sequence 16, Application US/11140224

Publication No. US20050227280A1

GENERAL INFORMATION:

APPLICANT: INCYTE GENOMICS, INC.

APPLICANT: DELEBEANE, Angelo M.; LAL, Preeti G.

APPLICANT: HAFALIA, April J.A.; PATTERSON, Chandra

APPLICANT: WALIA, Narinder K.; KEARNEY, Liam

APPLICANT: TRIBOULEY, Catherine M.; KHAN, Farrah A.

APPLICANT: YAO, Monique G.; BAUGHN, Mariah R.

APPLICANT: AZIMZAI, Yalda; ELLIOTT, Vicki S.

APPLICANT: NGUYEN, Dannel B.; GANDHI, Ameena R.

APPLICANT: YANG, Junning; HERNANDEZ, Roberto

APPLICANT: POLICKY, Jennifer L.; LU, Dyrung Aina M.

APPLICANT: REDDY, Roopa M.; YUE, Henry

APPLICANT: TANG, Y. Tom

TITLE OF INVENTION: PROTEASES

FILE REFERENCE: PI-0085 USN

CURRENT APPLICATION NUMBER: US/11/140,224

CURRENT FILING DATE: 2005-05-31

PRIOR APPLICATION NUMBER: US/10/275,505

PRIOR FILING DATE: 2002-11-04

PRIOR APPLICATION NUMBER: PCT/US01/14651

PRIOR FILING DATE: 2001-05-04

PRIOR APPLICATION NUMBER: 60/209,402

PRIOR FILING DATE: 2000-06-01

PRIOR APPLICATION NUMBER: 60/207,477

PRIOR FILING DATE: 2000-05-25

PRIOR APPLICATION NUMBER: 60/205,803

PRIOR FILING DATE: 2000-05-17

PRIOR APPLICATION NUMBER: 60/203,566

PRIOR FILING DATE: 2000-05-11

PRIOR APPLICATION NUMBER: 60/202,082

PRIOR FILING DATE: 2000-05-04

NUMBER OF SEQ ID NOS: 28

SOFTWARE: PERL Program

SEQ ID NO 16

LENGTH: 2510

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc feature

OTHER INFORMATION: Incyte ID No: 376067CB1

US-11-140-224-16

Qy	289	GATCATAGTTGATGAAGTCAGAAAGGCTGTATATTTTGAAGGCACCAAGACTCCCTTTT	348
Db	1410	GATCCAAGTTGATGAAGTCAGAAAGGCTGTATATTTTGAAGGCACCAAGACTCCCTTTT	1469
Qy	349	AGAGCATCACCTGTACGTAGTACGTAAATCTCTGGAGAGGTGACAAAGGCTGACTGA	408
Db	1470	AGAGCATCACCTGTACGTAGTACGTAAATCTCTGGAGAGGTGACAAAGGCTGACTGA	1529
Qy	409	CCGTGGCTACTCACATTTCTTCTGTCATCACTGACAGTCTGACTCTTTTATTAAGTAAGTA	468
Db	1530	CCGTGGCTACTCACATTTCTTCTGTCATCACTGACAGTCTGACTCTTTTATTAAGTAAGTA	1589
Qy	469	TAGTAACACAGAAAGTCCACACTGTGTCTTCAAGCTATCAAGTCTCTGAAGATGA	528
Db	1590	TAGTAACACAGAAAGTCCACACTGTGTCTTCAAGCTATCAAGTCTCTGAAGATGA	1649
Qy	529	CCCAACTTGCACAAAGGAATTTTGGGCCACCATTTTGGATTTCAGCAGGTCCTCTTCC	588
Db	1650	CCCAACTTGCACAAAGGAATTTTGGGCCACCATTTTGGATTTCAGCAGGTCCTCTTCC	1709
Qy	589	TGACTATACCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGATTTCACATTGTATGG	648
Db	1710	TGACTATACCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGATTTCACATTGTATGG	1769
Qy	649	GATGCTCTACAAGCCTCATGATCTACAGCCTGGAAGAAATATCTCTACTGTGCTGTTCAT	708
Db	1770	GATGCTCTACAAGCCTCATGATCTACAGCCTGGAAGAAATATCTCTACTGTGCTGTTCAT	1829
Qy	709	ATATGGTGGTCTCAGGTGCAAGTTCGTAATATCGTTTAAAGGAGTCAAGTATTTCCG	768
Db	1830	ATATGGTGGTCTCAGGTGCAAGTTCGTAATATCGTTTAAAGGAGTCAAGTATTTCCG	1889
Qy	769	CTTGAATACCTTAGCCTCTTAGGTTATGTGTTTGTAGTGATAGACAAACAGGGGATCCTG	828
Db	1890	CTTGAATACCTTAGCCTCTTAGGTTATGTGTTTGTAGTGATAGACAAACAGGGGATCCTG	1949
Qy	829	TCACCGAGGCTTAAATTTGAAGGCGCTTTAAATATATAAATGGGTCAAATAG	881
Db	1950	TCACCGAGGCTTAAATTTGAAGGCGCTTTAAATATATAAATGGGTCAAATAG	2002

Search completed: May 4, 2006, 11:21:59

Job time : 964.428 secs

Query Match

Best Local Similarity

Matches 586; Conservative

53.7%; Score 581.8; DB 10; Length 2510;

Pred. No. 4.2e-158;

0; Mismatches 7; Indels 0; Gaps 0;

THIS PAGE BLANK (USPTO)

GenCore version 5.1.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 4, 2006, 03:49:39 ; Search time 1479.29 Seconds
(without alignments)
2981.544 Million cell updates/sec

Title: US-10-825-632-8

Perfect score: 1083

Sequence: 1 ggaagaagatgccagatcag.....tatgatacagatacacgga 1083

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 9306428 seqs, 2036268586 residues

Total number of hits satisfying chosen parameters: 18612856

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.New.*
1: /SID55/ptodata/2/pubpna/US08_NEW_PUB.seq1.*
2: /SID55/ptodata/2/pubpna/US06_NEW_PUB.seq.*
3: /SID55/ptodata/2/pubpna/US07_NEW_PUB.seq.*
4: /SID55/ptodata/2/pubpna/US08_NEW_PUB.seq.*
5: /SID55/ptodata/2/pubpna/US09_NEW_PUB.seq.*
6: /SID55/ptodata/2/pubpna/US09_NEW_PUB.seq1.*
7: /SID55/ptodata/2/pubpna/US09_NEW_PUB.seq2.*
8: /SID55/ptodata/2/pubpna/US10_NEW_PUB.seq.*
9: /SID55/ptodata/2/pubpna/US10_NEW_PUB.seq1.*
10: /SID55/ptodata/2/pubpna/US10_NEW_PUB.seq2.*
11: /SID55/ptodata/2/pubpna/US10_NEW_PUB.seq3.*
12: /SID55/ptodata/2/pubpna/US10_NEW_PUB.seq4.*
13: /SID55/ptodata/2/pubpna/US11_NEW_PUB.seq.*
14: /SID55/ptodata/2/pubpna/US11_NEW_PUB.seq1.*
15: /SID55/ptodata/2/pubpna/US11_NEW_PUB.seq2.*
16: /SID55/ptodata/2/pubpna/US11_NEW_PUB.seq3.*
17: /SID55/ptodata/2/pubpna/US11_NEW_PUB.seq4.*
18: /SID55/ptodata/2/pubpna/US11_NEW_PUB.seq5.*
19: /SID55/ptodata/2/pubpna/US60_NEW_PUB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	791.2	73.1	2649	17	US-11-151-601-21
2	791.2	73.1	3143	17	US-11-151-601-19
3	118	10.9	1346	7	US-09-925-065A-669313
4	68.	6.3	2238	18	US-11-079-463-1186
5	59.6	5.5	2217	18	US-11-208-288-3
6	59.6	5.5	2301	10	US-10-522-789-1
7	59.6	5.5	3332	18	US-11-208-288-1
8	59.6	5.5	3407	11	US-10-501-035-34
9	58.4	5.4	4852	17	US-11-136-527-2130
10	54.8	5.1	1884	18	US-11-079-463-2536
11	54.8	5.1	2283	18	US-11-208-288-5
12	52.4	4.8	535	18	US-11-226-869-428
13	52.4	4.8	2788	9	US-10-505-928-476
14	52.4	4.8	2814	17	US-11-186-284-54

Sequence 158, App
Sequence 54125, A
Sequence 54125, A
Sequence 7940, Ap
Sequence 621349, A
Sequence 13487, A
Sequence 28205, A
Sequence 129442, A
Sequence 742851, A
Sequence 92266, A
Sequence 193508, A
Sequence 806917, A
Sequence 37, Appl
Sequence 83687, A
Sequence 14927, A
Sequence 798336, A
Sequence 515055, A
Sequence 457124, A
Sequence 797098, A
Sequence 718581, A
Sequence 718582, A
Sequence 67861, A
Sequence 401937, A
Sequence 469031, A
Sequence 1082440, A
Sequence 13396, A
Sequence 546889, A
Sequence 524839, A

ALIGNMENTS

RESULT 1

US-11-151-601-21
; Sequence 21, Application US/11151601
; Publication No. US20060003413A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Olandt, Peter J.
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Curtis, Rory A. J.
; APPLICANT: Williamson, Mark
; APPLICANT: Welch, Nadine
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASE, PHOSPHATASE,
; FILE REFERENCE: MPI00-054P1RCP10MIDVIM
; CURRENT APPLICATION NUMBER: US/11/151,601
; CURRENT FILING DATE: 2005-06-13
; PRIOR APPLICATION NUMBER: US 10/170,789
; PRIOR FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: US 09/797,039
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: PCT/US01/06525
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/186,061
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 09/882,166
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: PCT/US01/19269
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/212,078
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 09/934,406
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: PCT/US01/26052
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/226,740
; PRIOR FILING DATE: 2000-08-21

QY 640 ATTGTATGGGATGCTCTACAGCCTCATGATCTACAGCCTGGAAGAAATATCCTACTGT 699
Db 2115 ATTGTATGGGATGCTCTACAGCCTCATGATCTACAGCCTGGAAGAAATATCCTACTGT 2174
QY 700 GCTGTTCATATATGTTGCTCTCAGGTGCGAGTTGGTGAATAATCGGTTTAAAGGAGTCAA 759
Db 2175 GCTGTTCATATATGTTGCTCTCAGGTGCGAGTTGGTGAATAATCGGTTTAAAGGAGTCAA 2234
QY 760 GTATTTCCGCTTGAATACCTTAGCCTCTTAGGTTATGTTGTTAGTGTAGTATAGACAAACAG 819
Db 2235 GTATTTCCGCTTGAATACCTTAGCCTCTTAGGTTATGTTGTTAGTGTAGTATAGACAAACAG 2294
QY 820 GGGATCCGTGTCACCGAGGCTTAAATTTGAAGGCGCTTTAAATATAAAATGGGTCAAAT 879
Db 2295 GGGATCCGTGTCACCGAGGCTTAAATTTGAAGGCGCTTTAAATATAAAATGGGTCAAAT 2354
QY 880 AGAAATTGACGATCAGGTGGAAGGACTCCCAATATCTAGCTTCTCGATATGATTTCAATTGA 939
Db 2355 AGAAATTGACGATCAGGTGGAAGGACTCCCAATATCTAGCTTCTCGATATGATTTCAATTGA 2414
QY 940 CTTAGATCGTGGGACATCCACGCGTGGTCTCTATGAGGATACCTCTCCCTGATGGCAAT 999
Db 2415 CTTAGATCGTGGGACATCCACGCGTGGTCTCTATGAGGATACCTCTCCCTGATGGCAAT 2474
QY 1000 AATGACAGGTCAGATATCTTTACGGGTGCTATGTTGCGGGCCCAAGTCACTCTGTGGAT 1059
Db 2475 AATGACAGGTCAGATATCTTTACGGGTGCTATGTTGCGGGCCCAAGTCACTCTGTGGAT 2534
QY 1060 CTTCTATGATACAGGATACACGGA 1083
Db 2535 CTTCTATGATACAGGATACACGGA 2558

RESULT 3

US-09-925-065A-669313/c
; Sequence 669313, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wahg, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 669313
; LENGTH: 1346
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-669313

Query Match 10.9%; Score 118; DB 7; Length 1346;
Best Local Similarity 100.0%; Pred. No. 5.4e-22;
Matches 118; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 576 CAGGTCTCTTCCGACTATCTCTCCAGAAATTTCTCTTTTGAAGTACTACTGGAT 635
Db 118 CAGGTCTCTTCCGACTATCTCTCTCCAGAAATTTCTCTTTTGAAGTACTACTGGAT 59
QY 636 TTACATGTATGGATGCTCTACAGCCTCATGATCTACAGCCTGGAAGAAATATCC 693

Db 58 TTACATGTATGGGATGCTCTACAGCCTCATGATCTACAGCCTGGAAGAAATATCC 1

RESULT 4

US-11-079-463-1186
; Sequence 1186, Application US/11079463
; Publication No. US20060073161A1
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTEROIDES FRA
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: PATH00-03DIV2
; CURRENT APPLICATION NUMBER: US/11/079,463
; CURRENT FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/128,705
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: US 09/540,209
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 10444
; SEQ ID NO 1186
; LENGTH: 2238
; TYPE: DNA
; ORGANISM: B.fragilis
US-11-079-463-1186

Query Match 6.3%; Score 68; DB 18; Length 2238;
Best Local Similarity 49.6%; Pred. No. 4.3e-08;
Matches 238; Conservative 0; Mismatches 230; Indels 12; Gaps 2;

QY 604 AGAAATTTCTCTTTGAAAGTACTACTGGATTTACATGTTATGGGATGCTCTACAGCC 663
Db 1491 AGAATTTCTTACGTTTCAAAACACAGAGGAGTGCATCTGAACGGCTGGATGATGAAC 1550
QY 664 TCATGATCTACAGCTGGAAGAAATATCTACTGTGCTGCTTCATATATGGTGTCTCA 723
Db 1551 GGTCAATTTTCGATCTCGCAACGTTATCCGGTACTGATGTTCCAGTATAGCGGTCGGG 1610
QY 724 GGTGCAAGTTGTTGAATATCGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCCTAGC 783
Db 1611 TTCGCAACAGGTTCTGGACAAATGGGGAATCAGTTGGGAACCTCATG-----GC 1661
QY 784 CTCTCTAGGTTATGTTGTTAGTGAACACAGGGGATCTGTCAACCGAGGGCTTAA 843
Db 1662 GAGCCTCGGTTACGTTGTTAGTGTGTTAGTGTGCGGCACAGTGGCGGTGGCAGTGA 1721
QY 844 ATTTGAAGCGCCTTTAAATATAAAATGAGTCAAAATAGAAATTTGACGATCAGTGAAGG 903
Db 1722 ATTCAGAAATGCACCTACCTGAACTGGTGTAAAGAGCTTAAAGACCCAGGTGGAAGC 1781
QY 904 ACTCCAATATCTAGCTTCTCGATATGATTTTCATTTAGTCTGTGTTGGGATCCACGG 963
Db 1782 TGCCAAATATCTGGGTGGACTGTCCTTATGTG---GACAAAGGACGTTATTTGGTATCTGGGG 1838
QY 964 CTGGTCTTATGGAGATACCTCTCCCTGATGGCATTAATGACAGAGTCAAGATCTTTCAG 1023
Db 1839 ATGGAGTTTCGGCGGATATATGACCAATCATGATGATGAGCGAAGGTACACCCGTGTTAA 1898
QY 1024 GGTTCCTATTCTGGGGCCCGAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGA 1083
Db 1899 AGCCGGAGTTCTGTGGCGGACCTACAGACTGGAATATTACGATACAGTATATACCGA 1958

RESULT 5

US-11-208-288-3
; Sequence 3, Application US/11208288
; Publication No. US20060051366A1
; GENERAL INFORMATION:
; APPLICANT: CHANG, Chiwen
; TITLE OF INVENTION: USE OF SOLUBLE CD26 AS INHIBITOR OF
; TITLE OF INVENTION: ANGIOGENESIS AND INFLAMMATION
; FILE REFERENCE: 39533-0001
; CURRENT APPLICATION NUMBER: US/11/208,288
; CURRENT FILING DATE: 2005-08-18

; PRIOR APPLICATION NUMBER: US 60/605,013
; PRIOR FILING DATE: 2004-08-26
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 2217
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-208-288-3

Query Match 5.5%; Score 59.6; DB 18; Length 2217;
Best Local Similarity 53.6%; Pred. No. 9e-06;
Matches 147; Conservative 0; Mismatches 124; Indels 3; Gaps 1;

QY 810 TAGACAACAGGGGATCCTGTCCACCGAGGGCTTAAATTTGAAGGCCCTTTAAATATATAAAA 869
DB 1649 TTGATGCGACAGGAAGTGGTTACCAAGGAGATAAGATCATGCAATCAACAGAAGAC 1708

QY 870 TGGGTCAAATAGAAATGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATG 929
DB 1709 TGGGAACATTGAGATTGAAGATCAAAATTGAGCAGCCAGA---CAATTTTCAAAAATGG 1765

QY 930 ATTTTCATTGACTTAGATCGTGGGCATCCAGGCTGGTCTCTATGGAGATACCTCTCCC 989
DB 1766 GATTTGTGGACAACAACGAATTCGAATTTGGGGCTGGTCATATGAGGGTACGTAACT 1825

QY 990 TGATGGCAATTAATCCAGAGTCCAGATATCTTCAGGGTTGCTATTGCTGGGGCCCCAGTCA 1049
DB 1826 CAATGGTCTCGGATCCGGGAAGTGGCGGTGTTCAAGTGTGGAATACCGGTGGCGCTGTAT 1885

QY 1050 CTCGTGTGATCTTCTATGATACAGGATACACGA 1083
DB 1886 CCCGGTGGGAGTACTATGACTCATGTGTACACAGA 1919

RESULT 6
US-10-522-789-1
; Sequence 1, Application US/10522789
; Publication No. US20050260732A1
; GENERAL INFORMATION:
; APPLICANT: TANABE SEIYAKU CO., LTD.
; TITLE OF INVENTION: Three-dimensional structure of dipeptidyl peptidase IV
; FILE REFERENCE: 03-039-PCT
; CURRENT APPLICATION NUMBER: US/10/522,789
; PRIOR FILING DATE: 2005-01-28
; PRIOR APPLICATION NUMBER: US 60/398,761
; PRIOR FILING DATE: 2002-07-29
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 2301
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CD8
; LOCATION: (1)..(2301)
; OTHER INFORMATION:
US-10-522-789-1

Query Match 5.5%; Score 59.6; DB 10; Length 2301;
Best Local Similarity 53.6%; Pred. No. 9.1e-06;
Matches 147; Conservative 0; Mismatches 124; Indels 3; Gaps 1;

QY 810 TAGACAACAGGGGATCCTGTCCACCGAGGGCTTAAATTTGAAGGCCCTTTAAATATATAAAA 869
DB 1733 TTGATGCGACAGGAAGTGGTTACCAAGGAGATAAGATCATGCAATCAACAGAAGAC 1792

QY 870 TGGGTCAAATAGAAATGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATG 929
DB 1793 TGGGAACATTGAAAGTTGAAGATCAAAATTGAAGCAGCCAGA---CAATTTTCAAAAATGG 1849

QY 930 ATTTTCATTGACTTAGATCGTGGGCATCCAGGCTGGTCTCTATGGAGATACCTCTCCC 989
DB 1886 CCCGGTGGGAGTACTATGACTCATGTGTACACAGA 1919

DB 1850 GATTTGTGGACAACAACGAATTCGAATTTGGGGCTGGTCATATGAGGGTACGTAACT 1909
QY 990 TGATGGCAATTAATGACAGAGTCCAGATATCTTCAGGGTTCTCTATGCTGGGGCCCCAGTCA 1049
DB 1910 CAATGGTCTCGGATCGGGAAGTGGCGGTGTTCAAGTGTGGAATAGCCGTGGCGCTGTAT 1969

QY 1050 CTCGTGTGATCTTCTATGATACAGGATACACGA 1083
DB 1970 CCCGGTGGGAGTACTATGACTCATGTGTACACAGA 2003

RESULT 7
US-11-208-288-1
; Sequence 1, Application US/11208288
; Publication No. US20060051366A1
; GENERAL INFORMATION:
; APPLICANT: CHANG, Chiwen
; TITLE OF INVENTION: USE OF SOLUBLE CD26 AS INHIBITOR OF
; FILE REFERENCE: 39533-0001
; CURRENT APPLICATION NUMBER: US/11/208,288
; CURRENT FILING DATE: 2005-08-18
; PRIOR APPLICATION NUMBER: US 60/605,013
; PRIOR FILING DATE: 2004-08-26
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3332
; TYPE: DNA
; ORGANISM: homo sapiens
US-11-208-288-1

Query Match 5.5%; Score 59.6; DB 18; Length 3332;
Best Local Similarity 53.6%; Pred. No. 1.1e-05;
Matches 147; Conservative 0; Mismatches 124; Indels 3; Gaps 1;

QY 810 TAGACAACAGGGGATCCTGTCCACCGAGGGCTTAAATTTGAAGGCCCTTTAAATATATAAAA 869
DB 1733 TTGATGCGACAGGAAGTGGTTACCAAGGAGATAAGATCATGCAATCAACAGAAGAC 1792

QY 870 TGGGTCAAATAGAAATGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATG 929
DB 1793 TGGGAACATTGAAAGTTGAAGATCAAAATTGAAGCAGCCAGA---CAATTTTCAAAAATGG 1849

QY 930 ATTTTCATTGACTTAGATCGTGGGCATCCAGGCTGGTCTCTATGGAGGATACCTCTCCC 989
DB 1850 GATTTGTGGACAACAACGAATTCGAATTTGGGGCTGGTCATATGAGGGTACGTAACT 1909

QY 990 TGATGGCAATTAATGACAGAGTCCAGATATCTTCAGGGTTGCTATTGCTGGGGCCCCAGTCA 1049
DB 1910 CAATGGTCTCGGATCGGGAAGTGGCGGTGTTCAAGTGTGGAATAGCCGTGGCGCTGTAT 1969

QY 1050 CTCGTGTGATCTTCTATGATACAGGATACACGA 1083
DB 1970 CCCGGTGGGAGTACTATGACTCATGTGTACACAGA 2003

RESULT 8
US-10-501-035-34
; Sequence 34, Application US/10501035
; Publication No. US20060046249A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: IDENTIFICATION OF POLYNUCLEOTIDES AND POLYPEPTIDE FOR PREDICTING
; TITLE OF INVENTION: ACTIVITY OF COMPOUNDS THAT INTERACT WITH PROTEIN TYROSINE KINASE.
; TITLE OF INVENTION: AND/OR PROTEIN TYROSINE KINASE PATHWAYS
; FILE REFERENCE: D0185 PCT
; CURRENT APPLICATION NUMBER: US/10/501,035
; CURRENT FILING DATE: 2004-07-09
; PRIOR APPLICATION NUMBER: US 60/350,061
; PRIOR FILING DATE: 2002-01-18
; NUMBER OF SEQ ID NOS: 795
; SOFTWARE: PatentIn version 3.2

Db 2049 TGGGCGCGGTGTACAGCGTGGAGTACTATGACTCAGTATACACAGA 2094

RESULT 10
US-11-079-463-2536
; Sequence 2536, Application US/11079463
; Publication No. US20060073161A1
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTERIOIDES FR
; FILE REFERENCE: PATH00-03DIV2
; CURRENT APPLICATION NUMBER: US/11/079,463
; CURRENT FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/128,705
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: US 09/540,209
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 10444
; SEQ ID NO 2536
; LENGTH: 1884
; TYPE: DNA
; ORGANISM: B.fragilis
US-11-079-463-2536

Query Match 5.1%; Score 54.8; DB 18; Length 1884;
Best Local Similarity 47.6%; Pred. No. 0.00018;
Matches 195; Conservative 0; Mismatches 212; Indels 3; Gaps 1;
Qy 659 AGCCCTCATGCTACAGCGTGGAGAAATATCTACTGTGCTGTTCATATATGTTGGT 718
Db 1186 AAACCGCGCGATTTTCGATCCGAAACAAGAAATACCCGGCTATCGTATAGTATACGGCGT 1245
Qy 719 CCTCAGGTGAGTGTGTGAATATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACC 778
Db 1246 CCTCAGGTGAGTGTGTGAATATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACC 1305
Qy 779 CTAGCTCTCTAGTTTATGTTGTAGTATAGACAAACAGGGATCTGTCCACCGAGGG 838
Db 1306 ATGCCCAACAAAGTTTACATCATGTTTACCGTAGACGCGTGGAGCAATCGCGGA 1365
Qy 839 CTAAATTTGAAGCGCCCTTTAAATATATAAATAGGCTCAATAGAAATGACGATCAGTG 898
Db 1366 CTCGATTTGAGAATGTTACTTTCCGCCAGTTGGGAATCGAAGAGGAGACAGCGTG 1425
Qy 899 GAAGGACTCAATATCTAGCTTCTCGATATGATTTTCAATGACTTAGATCGTGTGGCATC 958
Db 1426 AAAGGAACCAATTTCTGAAAGCCCTCCCTACGTCGATGGGA---ACCGTATCGGATA 1482
Qy 959 CACGCTGTCTTATGAGGATACCTTCTCTGTATGGCAATTAATGACAGGTCAGATATC 1018
Db 1483 CACGCTGTGAGCTTTGGCGGTACATGACCACTGCCCTCTCTCTGCTATCCCGAGATA 1542
Qy 1019 TTCAGGTTGCTATTTGCTGGGGCCCGAGTCACTCTGTGGATCTTCTATGA 1068
Db 1543 TTTAAAGTGGCGTGGCGCGGCTCCGCTCAATGACTGGGGTATTACGA 1592

RESULT 11
US-11-208-288-5
; Sequence 5, Application US/11208288
; Publication No. US20060051366A1
; GENERAL INFORMATION:
; APPLICANT: CHANG, Chiwen
; TITLE OF INVENTION: USE OF SOLUBLE CD26 AS INHIBITOR OF
; FILE REFERENCE: 39533-0001
; CURRENT APPLICATION NUMBER: US/11/208,288
; CURRENT FILING DATE: 2005-08-18
; PRIOR APPLICATION NUMBER: US 60/605,013
; PRIOR FILING DATE: 2004-08-26
; NUMBER OF SEQ ID NOS: 8

Query Match 5.1%; Score 54.8; DB 18; Length 1884;
Best Local Similarity 47.6%; Pred. No. 0.00018;
Matches 195; Conservative 0; Mismatches 212; Indels 3; Gaps 1;
Qy 659 AGCCCTCATGCTACAGCGTGGAGAAATATCTACTGTGCTGTTCATATATGTTGGT 718
Db 1186 AAACCGCGCGATTTTCGATCCGAAACAAGAAATACCCGGCTATCGTATAGTATACGGCGT 1245
Qy 719 CCTCAGGTGAGTGTGTGAATATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACC 778
Db 1246 CCTCAGGTGAGTGTGTGAATATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACC 1305
Qy 779 CTAGCTCTCTAGTTTATGTTGTAGTATAGACAAACAGGGATCTGTCCACCGAGGG 838
Db 1306 ATGCCCAACAAAGTTTACATCATGTTTACCGTAGACGCGTGGAGCAATCGCGGA 1365
Qy 839 CTAAATTTGAAGCGCCCTTTAAATATATAAATAGGCTCAATAGAAATGACGATCAGTG 898
Db 1366 CTCGATTTGAGAATGTTACTTTCCGCCAGTTGGGAATCGAAGAGGAGACAGCGTG 1425
Qy 899 GAAGGACTCAATATCTAGCTTCTCGATATGATTTTCAATGACTTAGATCGTGTGGCATC 958
Db 1426 AAAGGAACCAATTTCTGAAAGCCCTCCCTACGTCGATGGGA---ACCGTATCGGATA 1482
Qy 959 CACGCTGTCTTATGAGGATACCTTCTCTGTATGGCAATTAATGACAGGTCAGATATC 1018
Db 1483 CACGCTGTGAGCTTTGGCGGTACATGACCACTGCCCTCTCTCTGCTATCCCGAGATA 1542
Qy 1019 TTCAGGTTGCTATTTGCTGGGGCCCGAGTCACTCTGTGGATCTTCTATGA 1068
Db 1543 TTTAAAGTGGCGTGGCGCGGCTCCGCTCAATGACTGGGGTATTACGA 1592

RESULT 12
US-11-208-288-5
; Sequence 5, Application US/11208288
; Publication No. US20060051366A1
; GENERAL INFORMATION:
; APPLICANT: CHANG, Chiwen
; TITLE OF INVENTION: USE OF SOLUBLE CD26 AS INHIBITOR OF
; FILE REFERENCE: 39533-0001
; CURRENT APPLICATION NUMBER: US/11/208,288
; CURRENT FILING DATE: 2005-08-18
; PRIOR APPLICATION NUMBER: US 60/605,013
; PRIOR FILING DATE: 2004-08-26
; NUMBER OF SEQ ID NOS: 8

Query Match 5.1%; Score 54.8; DB 18; Length 1884;
Best Local Similarity 47.6%; Pred. No. 0.00018;
Matches 195; Conservative 0; Mismatches 212; Indels 3; Gaps 1;
Qy 659 AGCCCTCATGCTACAGCGTGGAGAAATATCTACTGTGCTGTTCATATATGTTGGT 718
Db 1186 AAACCGCGCGATTTTCGATCCGAAACAAGAAATACCCGGCTATCGTATAGTATACGGCGT 1245
Qy 719 CCTCAGGTGAGTGTGTGAATATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACC 778
Db 1246 CCTCAGGTGAGTGTGTGAATATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACC 1305
Qy 779 CTAGCTCTCTAGTTTATGTTGTAGTATAGACAAACAGGGATCTGTCCACCGAGGG 838
Db 1306 ATGCCCAACAAAGTTTACATCATGTTTACCGTAGACGCGTGGAGCAATCGCGGA 1365
Qy 839 CTAAATTTGAAGCGCCCTTTAAATATATAAATAGGCTCAATAGAAATGACGATCAGTG 898
Db 1366 CTCGATTTGAGAATGTTACTTTCCGCCAGTTGGGAATCGAAGAGGAGACAGCGTG 1425
Qy 899 GAAGGACTCAATATCTAGCTTCTCGATATGATTTTCAATGACTTAGATCGTGTGGCATC 958
Db 1426 AAAGGAACCAATTTCTGAAAGCCCTCCCTACGTCGATGGGA---ACCGTATCGGATA 1482
Qy 959 CACGCTGTCTTATGAGGATACCTTCTCTGTATGGCAATTAATGACAGGTCAGATATC 1018
Db 1483 CACGCTGTGAGCTTTGGCGGTACATGACCACTGCCCTCTCTCTGCTATCCCGAGATA 1542
Qy 1019 TTCAGGTTGCTATTTGCTGGGGCCCGAGTCACTCTGTGGATCTTCTATGA 1068
Db 1543 TTTAAAGTGGCGTGGCGCGGCTCCGCTCAATGACTGGGGTATTACGA 1592

Db 2049 TGGGCGCGGTGTACAGCGTGGAGTACTATGACTCAGTATACACAGA 2094

RESULT 10
US-11-079-463-2536
; Sequence 2536, Application US/11079463
; Publication No. US20060073161A1
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTERIOIDES FR
; FILE REFERENCE: PATH00-03DIV2
; CURRENT APPLICATION NUMBER: US/11/079,463
; CURRENT FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/128,705
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: US 09/540,209
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 10444
; SEQ ID NO 2536
; LENGTH: 1884
; TYPE: DNA
; ORGANISM: B.fragilis
US-11-079-463-2536

Query Match 5.1%; Score 54.8; DB 18; Length 1884;
Best Local Similarity 47.6%; Pred. No. 0.00018;
Matches 195; Conservative 0; Mismatches 212; Indels 3; Gaps 1;
Qy 659 AGCCCTCATGCTACAGCGTGGAGAAATATCTACTGTGCTGTTCATATATGTTGGT 718
Db 1186 AAACCGCGCGATTTTCGATCCGAAACAAGAAATACCCGGCTATCGTATAGTATACGGCGT 1245
Qy 719 CCTCAGGTGAGTGTGTGAATATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACC 778
Db 1246 CCTCAGGTGAGTGTGTGAATATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACC 1305
Qy 779 CTAGCTCTCTAGTTTATGTTGTAGTATAGACAAACAGGGATCTGTCCACCGAGGG 838
Db 1306 ATGCCCAACAAAGTTTACATCATGTTTACCGTAGACGCGTGGAGCAATCGCGGA 1365
Qy 839 CTAAATTTGAAGCGCCCTTTAAATATATAAATAGGCTCAATAGAAATGACGATCAGTG 898
Db 1366 CTCGATTTGAGAATGTTACTTTCCGCCAGTTGGGAATCGAAGAGGAGACAGCGTG 1425
Qy 899 GAAGGACTCAATATCTAGCTTCTCGATATGATTTTCAATGACTTAGATCGTGTGGCATC 958
Db 1426 AAAGGAACCAATTTCTGAAAGCCCTCCCTACGTCGATGGGA---ACCGTATCGGATA 1482
Qy 959 CACGCTGTCTTATGAGGATACCTTCTCTGTATGGCAATTAATGACAGGTCAGATATC 1018
Db 1483 CACGCTGTGAGCTTTGGCGGTACATGACCACTGCCCTCTCTCTGCTATCCCGAGATA 1542
Qy 1019 TTCAGGTTGCTATTTGCTGGGGCCCGAGTCACTCTGTGGATCTTCTATGA 1068
Db 1543 TTTAAAGTGGCGTGGCGCGGCTCCGCTCAATGACTGGGGTATTACGA 1592

RESULT 11
US-11-208-288-5
; Sequence 5, Application US/11208288
; Publication No. US20060051366A1
; GENERAL INFORMATION:
; APPLICANT: CHANG, Chiwen
; TITLE OF INVENTION: USE OF SOLUBLE CD26 AS INHIBITOR OF
; FILE REFERENCE: 39533-0001
; CURRENT APPLICATION NUMBER: US/11/208,288
; CURRENT FILING DATE: 2005-08-18
; PRIOR APPLICATION NUMBER: US 60/605,013
; PRIOR FILING DATE: 2004-08-26
; NUMBER OF SEQ ID NOS: 8

Query Match 5.1%; Score 54.8; DB 18; Length 1884;
Best Local Similarity 47.6%; Pred. No. 0.00018;
Matches 195; Conservative 0; Mismatches 212; Indels 3; Gaps 1;
Qy 659 AGCCCTCATGCTACAGCGTGGAGAAATATCTACTGTGCTGTTCATATATGTTGGT 718
Db 1186 AAACCGCGCGATTTTCGATCCGAAACAAGAAATACCCGGCTATCGTATAGTATACGGCGT 1245
Qy 719 CCTCAGGTGAGTGTGTGAATATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACC 778
Db 1246 CCTCAGGTGAGTGTGTGAATATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACC 1305
Qy 779 CTAGCTCTCTAGTTTATGTTGTAGTATAGACAAACAGGGATCTGTCCACCGAGGG 838
Db 1306 ATGCCCAACAAAGTTTACATCATGTTTACCGTAGACGCGTGGAGCAATCGCGGA 1365
Qy 839 CTAAATTTGAAGCGCCCTTTAAATATATAAATAGGCTCAATAGAAATGACGATCAGTG 898
Db 1366 CTCGATTTGAGAATGTTACTTTCCGCCAGTTGGGAATCGAAGAGGAGACAGCGTG 1425
Qy 899 GAAGGACTCAATATCTAGCTTCTCGATATGATTTTCAATGACTTAGATCGTGTGGCATC 958
Db 1426 AAAGGAACCAATTTCTGAAAGCCCTCCCTACGTCGATGGGA---ACCGTATCGGATA 1482
Qy 959 CACGCTGTCTTATGAGGATACCTTCTCTGTATGGCAATTAATGACAGGTCAGATATC 1018
Db 1483 CACGCTGTGAGCTTTGGCGGTACATGACCACTGCCCTCTCTCTGCTATCCCGAGATA 1542
Qy 1019 TTCAGGTTGCTATTTGCTGGGGCCCGAGTCACTCTGTGGATCTTCTATGA 1068
Db 1543 TTTAAAGTGGCGTGGCGCGGCTCCGCTCAATGACTGGGGTATTACGA 1592

RESULT 12
US-11-208-288-5
; Sequence 5, Application US/11208288
; Publication No. US20060051366A1
; GENERAL INFORMATION:
; APPLICANT: CHANG, Chiwen
; TITLE OF INVENTION: USE OF SOLUBLE CD26 AS INHIBITOR OF
; FILE REFERENCE: 39533-0001
; CURRENT APPLICATION NUMBER: US/11/208,288
; CURRENT FILING DATE: 2005-08-18
; PRIOR APPLICATION NUMBER: US 60/605,013
; PRIOR FILING DATE: 2004-08-26
; NUMBER OF SEQ ID NOS: 8

Query Match 5.1%; Score 54.8; DB 18; Length 1884;
Best Local Similarity 47.6%; Pred. No. 0.00018;
Matches 195; Conservative 0; Mismatches 212; Indels 3; Gaps 1;
Qy 659 AGCCCTCATGCTACAGCGTGGAGAAATATCTACTGTGCTGTTCATATATGTTGGT 718
Db 1186 AAACCGCGCGATTTTCGATCCGAAACAAGAAATACCCGGCTATCGTATAGTATACGGCGT 1245
Qy 719 CCTCAGGTGAGTGTGTGAATATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACC 778
Db 1246 CCTCAGGTGAGTGTGTGAATATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACC 1305
Qy 779 CTAGCTCTCTAGTTTATGTTGTAGTATAGACAAACAGGGATCTGTCCACCGAGGG 838
Db 1306 ATGCCCAACAAAGTTTACATCATGTTTACCGTAGACGCGTGGAGCAATCGCGGA 1365
Qy 839 CTAAATTTGAAGCGCCCTTTAAATATATAAATAGGCTCAATAGAAATGACGATCAGTG 898
Db 1366 CTCGATTTGAGAATGTTACTTTCCGCCAGTTGGGAATCGAAGAGGAGACAGCGTG 1425
Qy 899 GAAGGACTCAATATCTAGCTTCTCGATATGATTTTCAATGACTTAGATCGTGTGGCATC 958
Db 1426 AAAGGAACCAATTTCTGAAAGCCCTCCCTACGTCGATGGGA---ACCGTATCGGATA 1482
Qy 959 CACGCTGTCTTATGAGGATACCTTCTCTGTATGGCAATTAATGACAGGTCAGATATC 1018
Db 1483 CACGCTGTGAGCTTTGGCGGTACATGACCACTGCCCTCTCTCTGCTATCCCGAGATA 1542
Qy 1019 TTCAGGTTGCTATTTGCTGGGGCCCGAGTCACTCTGTGGATCTTCTATGA 1068
Db 1543 TTTAAAGTGGCGTGGCGCGGCTCCGCTCAATGACTGGGGTATTACGA 1592

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 2283
; TYPE: DNA
; ORGANISM: Mus musculus
US-11-208-288-5

Query Match
Best Local Similarity 5.1%; Score 54.8; DB 18; Length 2283;
Matches 144; Conservative 0; Mismatches 127; Indels 3; Gaps 1;
QY 810 TAGACACAGGAGTCTCTGACCGAGGCTTAAATTTGAAGGCGCTTTAAATATAAAA 869
DB 1715 TTGACGCGACAGGAAGTGGTTACCAAGGAGATAAGCATCATCGATCAACAGAAT 1774
QY 870 TGGGTCAAAATAGAAATGACGATCGATGAGGACTCAATATCTAGCTTCTCGATG 929
DB 1775 TGGGAACACTGGAAGTTGAAGATCAATTTGAAGCAGCC--AGCAATTTGTAATGG 1831
QY 930 ATTTCATGACTTAGATCGTGGGCATCCAGCGCTGCTCTATGGAGGATACCTCTCC 989
DB 1832 GATTTGGATACCAAGCAGTTGCATTTGGGCTGGTCAATATGGAGGGTATGTAACT 1891
QY 990 TGATGCGATTAATGACAGGTGATATCTTCAGGTTGCTATGCTGGGCCCCAGTCA 1049
DB 1892 CAATGCTCTGGATCGGAAGTGGCGTGTTCAGTGGGAATAGCTGTGGCCTGTGT 1951
QY 1050 CTCTGGATCTTCTATGATGATGATGATGATGATGATGATGATGATGATGATG 1083
DB 1952 CACGGTGGAGTACTATGACTCAGTGTATACAGCA 1985

RESULT 12
US-11-226-869-428

; Sequence 428, Application US/11226869
; Publication No. US20060069054A1
; GENERAL INFORMATION:
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Sleath, Paul R.
; APPLICANT: Persing, David H.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.470C14
; CURRENT APPLICATION NUMBER: US/11/226.869
; CURRENT FILING DATE: 2005-09-13
; NUMBER OF SEQ ID NOS: 627
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 428
; LENGTH: 538
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-226-869-428

Query Match
Best Local Similarity 4.8%; Score 52.4; DB 18; Length 535;
Matches 150; Conservative 0; Mismatches 141; Indels 3; Gaps 1;
QY 790 AGGTATGTGTTCTAGTATGATGATGATGATGATGATGATGATGATGATGATG 849
DB 137 AGGATGTGATTCCTCTGTTGATGATGATGATGATGATGATGATGATGATGATG 196
QY 850 AGGCGCTTTAAATATAAATAGGTTGATGATGATGATGATGATGATGATGATGATG 909
DB 197 CTATGAGTGTATCGAAGCTGGTGTGTTATGAGTTGAAGACCAAGATTACAGCTG 256
QY 910 ATATCTAGCTTCTCGATGATGATGATGATGATGATGATGATGATGATGATGATG 969
DB 257 AAA---ATTGATGAAATGGTTTCAATGATGAAAGAAATAGCATAATGGGCTGTC 313
QY 970 CTATGGAGGATACCTCTCCCTGATGATGATGATGATGATGATGATGATGATGATG 1029
DB 314 CTATGGAGGATACGTTTCAATGATGATGATGATGATGATGATGATGATGATGATG 373

QY 1030 TATTGCTGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGA 1083
DB 374 TATAGCAGTGGCTCCAGTCTCCAGCTGGGAATATTACGCGTCTGTCTACACAGA 427

RESULT 13

US-10-505-928-476
; Sequence 476, Application US/10505928
; Publication No. US2006008532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/39178
; CURRENT APPLICATION NUMBER: US/10/505.928
; PRIOR FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363,019
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: PatentIn 3.2
; SEQ ID NO 476
; LENGTH: 2788
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-505-928-476

Query Match
Best Local Similarity 4.8%; Score 52.4; DB 9; Length 2788;
Matches 150; Conservative 0; Mismatches 141; Indels 3; Gaps 1;
QY 790 AGGTATGTGTTCTAGTATGATGATGATGATGATGATGATGATGATGATGATG 849
DB 1903 AGGATGTGATTCCTGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1962
QY 850 AGGCGCTTTAAATATAAATAGGTTGATGATGATGATGATGATGATGATGATGATG 909
DB 1963 CTATGAGTGTATCGAAGCTGGTGTGTTATGAAAGTTGAAGACCAAGATTAC 2022
QY 910 ATATCTAGCTTCTCGATGATGATGATGATGATGATGATGATGATGATGATGATG 969
DB 2023 AAA---ATTGATGAAATGGTGTTCATGATGAAAAAAGAAATAGCCATATGGG 2079
QY 970 CTATGGAGGATACCTCTCCCTGATGATGATGATGATGATGATGATGATGATGATG 1029
DB 2080 CTATGGAGGATACGTTTCAATGATGATGATGATGATGATGATGATGATGATGATG 2139
QY 1030 TATTGCTGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGA 1083
DB 2140 TATAGCAGTGGCTCCAGTCTCCAGTGGGAATATTACGCGTCTGTCTACACAGA 2193

RESULT 14

US-11-186-284-54
; Sequence 54, Application US/11186284
; Publication No. US20050266493A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Berger, Allison
; APPLICANT: Guillemette, Tracy L.
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Thibodeau, Stephen N.
; APPLICANT: Burgart, Lawrence J.
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; FILE REFERENCE: MP01-0392RNM
; CURRENT APPLICATION NUMBER: US/11/186.284
; CURRENT FILING DATE: 2005-07-21
; PRIOR APPLICATION NUMBER: US/10/301,822
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 60/339,971
; PRIOR FILING DATE: 2001-12-10

Query Match 4.8%; Score 52.4; DB 17; Length 2814;
Best Local Similarity 51.0%; Pred. No. 0.00097;
Matches 150; Conservative 0; Mismatches 141; Indels 3; Gaps 1;
QY 790 AGGTTATGTGGTTGTAGTATAGACAAACAGGGGATCCTGTCCACCGAGGGCTTAAATTTGA 849
Db 1903 AGGGATGGTCATTGGCTTGGTGGATGGTCGAGGAAACAGCTTTTCCAAAGGTGACAACTCCT 1962
QY 850 AGGGGCTTTAAATATATAAATGAGTCAAAATGACGATGAGGTGGAAGGACTCCA 909
Db 1963 CTATGCAGTGTATCGAAAGCTGGGTGTTTATGAAGTTGAAGACCAAGATTACAGCTGTGAG 2022
QY 910 ATATCTAGCTTCTCGATATGATTTCACTTATGATCGTGTGGGCATCCACGGCTGTC 969
Db 2023 AAA---ATTTCATAGAAATGGGTTTCATTGATGAAAAAAGAAATAGCCATATGGGGCTGTC 2079
QY 970 CTATGGAGGATACCTCTCCCTGATGGCATTATGCGAGGTCAGATATCTTCAGGGTTGC 1029
Db 2080 CTATGGAGGATACGTTTCATCCTGAGGCTTGCATCTGGAATCTGAACTGGTCTTTTCAAATGTGG 2139
QY 1030 TATTGCTGGGGCCCCCAGTCACCTCTGTGGATCTTCTATGATACAGGATACACGGA 1083
Db 2140 TATAGCAGTGGCTCCAGTCTCCAGCTGGGAATATTACGGCTCTGTCTACACAGA 2193

Search completed: May 4, 2006, 09:26:09
Job time : 1481.29 secs

PRIOR APPLICATION NUMBER: US 60/361,978
PRIOR FILING DATE: 2002-03-05
PRIOR APPLICATION NUMBER: US 60/381,988
PRIOR FILING DATE: 2002-05-20
NUMBER OF SEQ ID NOS: 228
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 54
LENGTH: 2814
TYPE: DNA
ORGANISM: Homo Sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (209)....(2491)
US-11-186-284-54

Query Match 4.8%; Score 52.4; DB 17; Length 2814;
Best Local Similarity 51.0%; Pred. No. 0.00097;
Matches 150; Conservative 0; Mismatches 141; Indels 3; Gaps 1;
QY 790 AGGTTATGTGGTTGTAGTATAGACAAACAGGGGATCCTGTCCACCGAGGGCTTAAATTTGA 849
Db 1903 AGGGATGGTCATTGGCTTGGTGGATGGTCGAGGAAACAGCTTTTCCAAAGGTGACAACTCCT 1962
QY 850 AGGGGCTTTAAATATATAAATGAGTCAAAATGACGATGAGGTGGAAGGACTCCA 909
Db 1963 CTATGCAGTGTATCGAAAGCTGGGTGTTTATGAAGTTGAAGACCAAGATTACAGCTGTGAG 2022
QY 910 ATATCTAGCTTCTCGATATGATTTCACTTATGATCGTGTGGGCATCCACGGCTGTC 969
Db 2023 AAA---ATTTCATAGAAATGGGTTTCATTGATGAAAAAAGAAATAGCCATATGGGGCTGTC 2079
QY 970 CTATGGAGGATACCTCTCCCTGATGGCATTATGCGAGGTCAGATATCTTCAGGGTTGC 1029
Db 2080 CTATGGAGGATACGTTTCATCCTGAGGCTTGCATCTGGAATCTGAACTGGTCTTTTCAAATGTGG 2139
QY 1030 TATTGCTGGGGCCCCCAGTCACCTCTGTGGATCTTCTATGATACAGGATACACGGA 1083
Db 2140 TATAGCAGTGGCTCCAGTCTCCAGCTGGGAATATTACGGCTCTGTCTACACAGA 2193

RESULT 15
US-11-245-147-168
Sequence 168, Application US/11245147
Publication No. US20060030541A1
GENERAL INFORMATION:
APPLICANT: GARCIA, TERESA
APPLICANT: ROMAN, ROMAN, SERGIO
APPLICANT: BARON, ROLAND
APPLICANT: CALL, KATHERINE
APPLICANT: THEILHABER, JOACHIM
APPLICANT: CONNOLLY, TIMOTHY
APPLICANT: JACKSON, AMANDA
APPLICANT: BUSHNELL, STEVEN
APPLICANT: RAWADI, GEORGES
TITLE OF INVENTION: GENES INVOLVED IN OSTEOGENESIS, AND METHODS OF USE
FILE REFERENCE: 37991-0023
CURRENT APPLICATION NUMBER: US/11/245,147
CURRENT FILING DATE: 2005-10-07
PRIOR APPLICATION NUMBER: PCT/IB02/02211
PRIOR FILING DATE: 2002-04-05
PRIOR APPLICATION NUMBER: 60/281,400
PRIOR FILING DATE: 2001-04-05
NUMBER OF SEQ ID NOS: 246
SOFTWARE: Patent In Ver. 3.2
SEQ ID NO 168
LENGTH: 2814
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: Homo sapiens fibroblast activation protein, alpha
OTHER INFORMATION: (FAP), mRNA
US-11-245-147-168

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame plus n2p model

Run on: May 2, 2006, 00:41:48 ; Search time 32.8822 Seconds
(without alignments)
2894.257 Million cell updates/sec

Title: US-10-825-632-8
Perfect score: 1938
Sequence: 1 ggaagaagatgccagatcag.....tatgatacaggatacacgga 1083

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 4886326

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ n2p.model -DEV=xlp
-Q=/abss/ABSSWEB.spool/US10825632/runat_01052006_105944_3214/app.query.fasta.1
-DB=A.Geneseq -QFWT=fastcan -SUFFIX=n2p.rag -MINMATCH=0.1 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFWT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abss02p
-USER=US10825632@cgn 1.1.605 @runat_01052006_105944_3214 -NCPU=6 -ICPU=3
-NO MWAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A.Geneseq 21.*

- 1: Geneseqp1980s.*
- 2: Geneseqp1990s.*
- 3: Geneseqp2000s.*
- 4: Geneseqp2001s.*
- 5: Geneseqp2002s.*
- 6: Geneseqp2003as.*
- 7: Geneseqp2003bs.*
- 8: Geneseqp2004s.*
- 9: Geneseqp2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1933	99.7	360	4 AAB47190	AAB47190 Human DPP
2	1836.5	94.8	824	6 ABU92030	ABU92030 Human pro
3	1836.5	94.8	882	4 AAB47187	AAB47187 Human DPP
4	1836.5	94.8	882	5 ABG61591	ABG61591 Human DPP
5	1836.5	94.8	882	5 AAG78415	AAG78415 Amino aci
6	1836.5	94.8	882	5 AAE24170	AAE24170 Human dip
7	1836.5	94.8	882	5 AAU74749	AAU74749 Human pro
8	1836.5	94.8	882	5 ADI17086	ADI17086 Human NOV
9	1836.5	94.8	882	6 ABU07720	ABU07720 Human ser

10	1756.5	90.6	883	5 ADI17085	ADI17085 Murine NO
11	1534	79.2	831	6 ABU92026	ABU92026 Human pro
12	1533.5	79.1	493	7 ADE78977	ADE78977 Human pro
13	1528	78.8	632	4 AAB93565	AAB93565 Human pro
14	1525.5	78.7	587	5 ADR41398	ADR41398 Human CD-
15	1495	77.1	580	5 AAE14337	AAE14337 Human pro
16	1466.5	75.7	746	8 ADI16334	ADI16334 Human pro
17	1254.5	64.7	724	5 ABB97362	ABB97362 Novel hum
18	1254.5	64.7	782	5 ABB97361	ABB97361 Novel hum
19	1226	63.3	690	5 ABG61594	ABG61594 Human DPR
20	1220.5	63.0	689	5 ADI17084	ADI17084 Human NOV
21	1220.5	63.0	830	5 AAE24171	AAE24171 Human dip
22	1220.5	63.0	863	5 ABG61592	ABG61592 Human DPP
23	1220.5	63.0	863	5 ADI17083	ADI17083 Human NOV
24	1220.5	63.0	892	5 ABG61602	ABG61602 Human DPR
25	1220.5	63.0	892	5 ABG61604	ABG61604 Human DPR
26	1220.5	63.0	892	5 ABB98134	ABB98134 Human PMM
27	1220.5	63.0	892	8 ADS10951	ADS10951 Human the
28	1220.5	63.0	969	5 AAE24168	AAE24168 Human dip
29	1219	62.9	516	6 ABU92029	ABU92029 Human pro
30	1213.5	62.6	755	8 ADQ67811	ADQ67811 Novel hum
31	1212.5	62.6	863	5 ADI16690	ADI16690 Human NOV
32	1212.5	62.6	863	5 ADI16688	ADI16688 Human NOV
33	1212.5	62.6	863	8 ADN42344	ADN42344 Human nov
34	1209	62.4	661	5 ABG61596	ABG61596 Murine di
35	1204.5	62.2	847	5 AAE23875	AAE23875 Alternati
36	1204.5	62.0	869	5 AAE24169	AAE24169 Human DPR
37	1201	62.0	879	5 ABG61607	ABG61607 Human DPR
38	1201	62.0	879	5 ABG61608	ABG61608 Human DPR
39	1201	62.0	879	8 ADS10953	ADS10953 Human the
40	1199.5	61.9	658	5 ABG61600	ABG61600 Human DPR
41	1196.5	61.7	863	8 ADN42342	ADN42342 Human nov
42	1183.5	61.1	683	4 AAM40510	AAM40510 Human pol
43	1183.5	61.1	737	4 AAM38724	AAM38724 Human pol
44	1183.5	61.1	873	8 ADS10479	ADS10479 Human the
45	1174.5	60.6	854	8 ADS10952	ADS10952 Human the

ALIGNMENTS

RESULT 1
AAB47190
ID : AAB47190 standard; protein; 360 AA.
XX
AC: AAB47190;
XX
DT 29-JUN-2001 (first entry)
XX
DE Human DPP8 244Glu-341Ile+515Val-776Thr.
XX
KW Human; dipeptidyl aminopeptidase; DPP8; prolyl oligopeptidase;
KW dipeptidyl peptidase; DPPIV; T cell; cleavage; diarrhoea;
KW growth hormone deficiency; Glucose level; mucosal regeneration;
KW non-insulin dependent diabetes mellitus; glucose intolerance;
KW immunosuppression.
XX
OS Homo sapiens.
XX
PN WO200119866-A1.
XX
PD 22-MAR-2001.
XX
PF 11-SEP-2000; 2000WO-AU001085.
XX
PR 10-SEP-1999; 99AU-00002762.
XX
PR 18-FEB-2000; 2000AU-00005709.
XX
PA (UNSY) UNIV SYDNEY.
XX
PI Abbott CA, Gorell MD;
XX
DR WPI; 2001-281520/29.
DR N-PSDB; AAC85697.

XX New human dipeptidyl aminopeptidase (DPP8) useful for cleaving
 PT substrates, identifying inhibitors of DPP8 catalytic activity which have
 PT therapeutic uses, and for detecting activated T cells.
 XX
 PS Claim 5; Page 74-75; 78pp; English.

XX The sequences given in ABA47188-90 represent fragments of human
 CC dipeptidyl aminopeptidase (DPP8). DPP8 has substrate specificity for H-
 CC Ala-Pro-PNA, H-Gly-Pro-PNA and H-Arg-Pro-PNA. Therefore, it is a prolyl
 CC oligopeptidase and a dipeptidyl peptidase, because it is capable of
 CC hydrolysing the peptide bond C-terminal to pro in each of these
 CC compounds. DPP8 is homologous with human DPP4V. DPP8 is useful for
 CC cleaving a substrate, and for detecting an activated T cell which
 CC involves measuring the level of DPP8 gene expression in a T cell. The
 CC level of DPP8 expression is detected by detecting the amount of DPP8 RNA
 CC in the cell. It is also useful for identifying a molecule capable of
 CC inhibiting the cleavage of the substrate by DPP8. Molecules identified as
 CC inhibiting DPP8 catalytic activity may be useful for treating diarrhoea,
 CC growth hormone deficiency, lowering glucose levels in non-insulin
 CC dependent diabetes mellitus and other disorders involving glucose
 CC intolerance, enhancing mucosal regeneration and as immunosuppressants

XX Sequence 360 AA;

Alignment Scores:

Pred. No.: 6,47e-216 Length: 360
 Score: 1933.00 Matches: 360
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 99.7% Indels: 0
 DB: Gaps: 4

US-10-825-632-8 (1-1083) x ABA47190 (1-360)

QY 2 GAAGAGATGCCAGATCAGCTGGAGTCCCTACCTTTGTTCTCCAGAGAAATTTGATAGA 61
 DB 1 GluGluAspAlaArgSerAlaGlyValAlaIamrPheValLeuGluGluPheAspArg 20
 QY 62 TATTCCTGCTATTTGGTGGTCTCAAAAGCTGAACATCCCACTGGTGGTAAATCTT 121
 DB 21 TyrSerGlyTyrTrpCysProLysAlaGluThrThrProSerGlyGlyLysIleLeu 40
 QY 122 AGAATTCATATGAAGAAATGATGATCTGAGTGGGAAATTTATTCATGTACATCCCT 181
 DB 41 ArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisValThrSerPro 60
 QY 182 ATTTGGAAACAAGGAGGCGAGATTCATTCGTTATCTCAAAACAGGTACAGCAATCTCT 241
 DB 61 MetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThrAlaAsnPro 80
 QY 242 AAAGTCACCTTTAAGATGTCAGAAATAATGATGATGCTGAAGGAGGATCATAGTTGAT 301
 DB 81 LysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIleIleValAsp 100
 QY 302 GAAGTCAGAGGCTGGTATATTTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTG 361
 DB 101 GluValArgAlaLeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeu 120
 QY 362 TAGCTAGTCAGTTACGTAATCTCGAGAGGTGACAGGCTGACCTGCGCTACTCA 421
 DB 121 TyrValValSerTyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSer 140
 QY 422 CATTCCTGTCATCAGTCAGCACGTGCTCTTTTATAGTAAGTATAGTACACAGAG 481
 DB 141 HisSerCysGlySerGlnHisCysAspPhePheIleSerLysTyrSerArgGlnLys 160
 QY 482 AATCCACATGTGTGCTTTCAGAGCTATCAAGTCTCAAGTATGACCCCACTTCCAAA 541
 DB 161 AsnProHisCysValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLys 180
 QY 542 ACAAGGAATTTGGCCACCACTTTTGATTCACAGGTCTCTTCTGACTATATCTCT 601

DB 181 ThrLysGluPheTrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrPro 200
 QY 602 CCAGAAATTTCTCTTTTGAAGATGATCTGAGATTTACATGATGGGATGCTCTACAG 661
 DB 201 ProGluIlePheSerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLys 220
 QY 662 CCTCATGATCTACAGCTGGAATAATATCTCTACTGTCTGTTCATATATGCTGCTCT 721
 DB 221 ProHisAspLeuGlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyPro 240
 QY 722 CAGGTGAGTGTGGTGAATAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCCCTA 781
 DB 241 GlnValGlnLeuValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeu 260
 QY 782 GCCTCTCTAGGTATGTGTTGTAGTATAGACAAACAGGGGATCTCTGTCACGAGGCTT 841
 DB 261 AlaSerLeuGlyTyrValValValIleAspAsnArgGlySerCysHisArgGlyLeu 280
 QY 842 AAATTTCAAGGCGCTTTAAATATAAATGGGTCAAAATAGAAATGACGATCAGGTGGA 901
 DB 281 LysPheGluGlyAlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGlu 300
 QY 902 GGACTCCAATATCTAGCTTCTCGATATGATTTTCATTTGACTTATAGATCGTGGCATCCAC 961
 DB 301 GlyLeuGlnTyrLeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHis 320
 QY 962 GGCCTGGTCCCTATGGAGGATACCTCTCCCTGATGGCATTAATCGAGGTCAGATATCTTC 1021
 DB 321 GlyTrpSerTyrGlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePhe 340
 QY 1022 AGGCTGCTATTCCTGGGCGCCAGTCACCTCTGTGATCTTCTATGATACAGGATACAG 1081
 DB 341 ArgValAlaIleAlaGlyAlaProValThrLeuThrIlePheTyrAspThrGlyTyrThr 360

RESULT 2
 ABUS2030
 ID ABUS2030 standard; protein; 824 AA.
 XX
 AC ABUS2030;
 DT 15-JUL-2003 (first entry)
 XX
 DE Human protein modification and maintenance molecule-10 (PMMW-10).
 XX
 KW Human; protein modification and maintenance molecule; PMMW; cancer;
 KW cell proliferation disorder; atherosclerosis; neurological disorder;
 KW epilepsy; Huntington's disease; stroke; immune disorder; allergy;
 KW inflammatory disorder; AIDS; developmental disorder; hypothyroidism;
 KW Cushing's syndrome; gastrointestinal disorder; epithelial disorder;
 KW infection; cytostatic; antiarteriosclerotic; anticonvulsant; nootropic;
 KW neuroprotective; cerebroprotective; anti-HIV; antiallergic; vulnery;
 KW antiinflammatory; thyromimetic.
 XX
 OS Homo sapiens.
 XX
 PN WO2003031939-A2.
 XX
 PD 17-APR-2003.
 XX
 PF 11-OCT-2002; 2002WO-US032850.
 XX
 PR 12-OCT-2001; 2001US-0329689P.
 PR 25-OCT-2001; 2001US-0335703P.
 PR 09-NOV-2001; 2001US-0346887P.
 PR 28-NOV-2001; 2001US-0334145P.
 PR 06-DEC-2001; 2001US-0337451P.
 PR 14-DEC-2001; 2001US-0340584P.
 XX
 FA (INCY-) INCYTE GENOMICS INC.
 XX
 PI Runkumar J, Gorvad AB, Baughn MR, Emerling BM, Yang J, Lee SY;
 PI Tran UK, Becha SD, Duggan BM, Lee EA, Griffin JA, Li JX;
 PI Sprague WW, Hafalia AJA, Chawla NK, Lehr-Mason PM, Kable AB, Yue H;

PI Marquis JP, Yao MG, Richardson TW, Tang TY, Jin P, Chien D;
PI Bhatia U, Burrill JD, Lee S, Blake JJ, Ho A, Zheng W;
XX
DR WPI; 2003-430274/40.
DR N-PSDB; ACA92425.
XX
PT New human protein modification and maintenance molecules (PMMM), useful
PT for diagnosing, treating and preventing diseases or conditions associated
PT with the aberrant PMMM expression e.g. cancer, atherosclerosis, or
PT infections.

XX Claim 1; Page 249-251; 311pp; English.

XX The present invention relates to the isolation of human protein
CC modification and maintenance molecules (PMMM), and the polynucleotide
CC sequences encoding them. A total of 40 PMMM polypeptides (designated PMMM
CC -1 to PMMM-40) are disclosed. The sequences of the invention are useful
CC for diagnosing a condition or disease associated with the expression of
CC PMMM in a subject, preparing a polyclonal or monoclonal antibody, and
CC generating an expression profile of a sample containing the
CC polynucleotides. The diseases or conditions associated with decreased
CC expression or overexpression of PMMM are cell proliferation disorders
CC (e.g. cancer, atherosclerosis), neurological disorders (e.g. epilepsy,
CC Huntington's disease, stroke), immune/inflammatory disorders, (e.g. AIDS,
CC allergies), developmental disorders (e.g. hypothyroidism, Cushing's
CC syndrome), gastrointestinal or epithelial disorders, and infections. The
CC PMMM polypeptides or their fragments are useful in screening compounds
CC for effectiveness as agonists or antagonists of the polypeptides, or in
CC altering the expression of the target polynucleotide and compounds that
CC specifically bind to, or modulate the activity of the polypeptide.
CC ABU92021-ABU92060 represent the human PMMM polypeptides of the invention

XX Sequence 824 AA;

Alignment Scores:

Pred. No.: 1, 7e-204 Length: 824
Score: 1836.50 Matches: 360
Percent Similarity: 67.5% Conservative: 0
Best Local Similarity: 67.5% Mismatches: 0
Query Match: 94.8% Indels: 173
DB: 6 Gaps: 1

US-10-825-632-8 (1-1083) x ABU92030 (1-824)

QY 2 GAGAGATGCCAGATCAGTCGAGTCGCTACCTTTGTTCTCCAGAGAAATTTGATAGA 61
DB 186 GluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGluPheAspArg 205
QY 62 TATTCTGGCTATTGGTGGTCCAAAGCTGAAACAACTCCCGAGTGGTGTAAATTCCT 121
DB 206 TyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGlyLysIleLeu 225
QY 122 AGAATTCATTAAGAATAATGATGATCTGAGGTGGAAATTTATCATGTTACATCCCT 181
DB 226 ArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleHisValThrSerPro 245
QY 182 ATGTGGAACAAGGAGGCGAGATTCATTCGGTTATCTTAAACAGGTFACAGCAATCCT 241
DB 246 MetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThrAlaAsnPro 265
QY 242 AAAGTCACTTTTAAGATGTGAGAAATAATGATTGATGCTGAGGAGGATCATATA 295
DB 266 LysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIleIleAspVal 285
QY 295 ----- 295
DB 286 IleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGluTyrIleAla 305
QY 295 ----- 295
DB 306 ArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAspArgSerGln 325
QY 295 ----- 295

DB 326 ThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGluAspVal 345
QY 295 ----- 295
DB 346 MetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeuIleIleTyr 365
QY 295 ----- 295
DB 366 GluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPheProGlnSer 385
QY 295 ----- 295
DB 386 HisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPheArgHisLeu 405
QY 295 ----- 295
DB 406 TyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGlyGlyLeuPro 425
QY 295 ----- 295
DB 426 AlaProSerAspPheLysCysProLysGluGluIleAlaIleThrSerGlyGluTrp 445
QY 296 ----- GTTGATGAAGTCAGAGGCTGTATAT 322
DB 446 GluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgGluLeuValTyr 465
QY 323 TTTGAAGCGACCAAGATCCCTTTAGACATCACCCTAGCTAGTACGTACGTACGTAAAT 382
DB 466 PheGluGlyThrLysAspSerProLeuGluHisLysLeuTyrValValSerTyrValAsn 485
QY 383 CCTGAGAGGTGACAGGCTGACTGACGCTGGCTACTCACATTTCTGTGTCATGTCAGCAG 442
DB 486 ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysLysIleSerGln 505
QY 443 CACTGTGACTTCTTTTATAAGTATAGTAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 502
DB 506 HisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeu 525
QY 503 TACAAGCTATCAAGTCTCTGAAGATGACCCCAACTGTCACAAACAAAGAAATTTGGGCCACC 562
DB 526 TyrLysLeuSerSerProGluAspProThrCysLysThrLysGluPheTrpAlaThr 545
QY 563 ATTTTGAATTCAGCAGGCTCTCTCTCTGACTACTACTCTCCAGAAATTTCTCTTTTGA 622
DB 546 IleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePheSerPheGlu 565
QY 623 AGTACTACTGATTTACATTTGATGGATGCTCTCAAGCCTCATGATCATCAGCCTGGA 682
DB 566 SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly 585
QY 683 AAGAAATATCTACTGTGCTGTTTCATATATGGTGGTCTCAGTGCAGTGGTGAATAT 742
DB 586 LysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeuValAsnAsn 605
QY 743 CGGTTTAAAGAGTCAAGTATTTCCGCTTGAATPACCTAGCCTCTCTAGGTTATGTGTT 802
DB 606 ArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrValVal 625
QY 803 GTAGTGATAGCAACAGGGGATCTCTCACCAGGGCTTAAATTTCAAGGCCCTTTAAA 862
DB 626 ValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLys 645
QY 863 TATAAATGGTCAAAATAGAAATTTGACGATCAGTGGTGAAGGACTCCAAATATCTAGCTTCT 922
DB 646 TyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyrLeuAlaSer 665
QY 923 CGATATGATTTTCAATGATTTAGTTCGTGGGCATCCACGGCTGTCCTATGAGGATAC 982
DB 666 ArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyrGlyGlyTyr 685
QY 983 CTCCTCCCTGATGGCATTAAATGACAGTCAAGTATCTTCAAGGGTTCCTATTCCTGGGCC 1042

Db 686 LeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAlaGlyAla 705
 QY 1043 CCAGTCACTCTGGGATCTCTATGATACAGATACAG 1081
 Db 706 ProValThrLeuTrpIlePheTyrAspThrGlyTyrThr 718

RESULT 3

AA047187
 ID AAB47187 standard; protein; 882 AA.

AC AAB47187;

DI 29-JUN-2001 (first entry)

DE Human DPP8.

XX Human; dipeptidyl aminopeptidase; DPP8; prolly oligopeptidase;
 KW dipeptidyl peptidase; DPP8; T cell; cleavage; diarrhoea;
 KW growth hormone deficiency; glucose level; mucosal regeneration;
 KW non-insulin dependent diabetes mellitus; glucose intolerance;
 KW immunosuppression.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Active-site 739
 FT /note= "Forms part of Ser-Asp-His catalytic triad"
 FT Active-site 817
 FT /note= "Forms part of Ser-Asp-His catalytic triad"
 FT Active-site 849
 FT /note= "Forms part of Ser-Asp-His catalytic triad"

XX WO200119866-A1.

XX 22-MAR-2001.

XX 11-SEP-2000; 2000WO-AU001085.

XX 10-SEP-1999; 99AU-00002762.

XX 18-FEB-2000; 2000AU-00005709.

XX (UNSY) UNIV SYDNEY.

XX Abbott CA, Gorell MD;

XX WPI; 2001-281520/29.

XX N-PSDB; AAC85694.

XX New human dipeptidyl aminopeptidase (DPP8) useful for cleaving
 PT substrates, identifying inhibitors of DPP8 catalytic activity which have
 PT therapeutic uses, and for detecting activated T cells.

XX Claim 1; Fig 2; 78pp; English.

XX This sequence represents human dipeptidyl aminopeptidase (DPP8). DPP8 has
 CC substrate specificity for H-Ala-Pro-pNA, H-Gly-Pro-pNA and H-Arg-Pro-pNA.
 CC Therefore, it is a prolly oligopeptidase and a dipeptidyl peptidase.
 CC because it is capable of hydrolysing the peptide bond C-terminal to pro
 CC in each of these compounds. DPP8 is homologous with human DPPIV. DPP8 is
 CC useful for cleaving a substrate and for detecting an activated T cell
 CC which involves measuring the level of DPP8 gene expression in a T cell.
 CC The level of DPP8 expression is detected by detecting the amount of DPP8
 CC RNA in the cell. It is also useful for identifying a molecule capable of
 CC inhibiting the cleavage of the substrate by DPP8. Molecules identified as
 CC inhibiting DPP8 catalytic activity may be useful for treating diarrhoea,
 CC growth hormone deficiency, lowering glucose levels in non-insulin
 CC dependent diabetes mellitus and other disorders involving glucose
 CC intolerance, enhancing mucosal regeneration and as immunosuppressants

XX Sequence 882 AA;

SQ Alignment Scores:

Pred. No.: 1.75e-204 Length: 882

Score: 1836.50 Matches: 360
 Percent Similarity: 67.5% Conservative: 0
 Best Local Similarity: 67.5% Mismatches: 0
 Query Match: 94.8% Indels: 173
 DB: 4 Gaps: 1
 US-10-825-632-8 (1-1083) x AAB47187 (1-882)
 QY 2 GAAGAAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAAGAAGAAATTTGATAGA 61
 Db 244 GluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGluPheAspArg 263
 QY 62 TATTCTGGCTATTGGTGTCTCCAAAAGCTGAAACAACTCCAGAGTGGTGGTAAAAATTCCT 121
 Db 264 TyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGlyLysIleLeu 283
 QY 122 AGAATTCCTATATGAAGAAATGATGATCTGAGGTGGAATATTATTCATGTTCATCCCT 181
 Db 284 ArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisValThrSerPro 303
 QY 182 ATGTTGAAACAAAGGAGGCGAGATTCATTCCCTTATCTTAAACACAGGTACAGCAATCCT 241
 Db 304 MetLeuGluThrArgArgAlaAspSerPheA-GTyrProLysThrGlyThrAlaAsnPro 323
 QY 242 AAAGTCATTTTAAGATCTCAGAAATTAATGATGATGCTGAAGGAAGGATCATATA 295
 Db 324 LysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIleIleAspVal 343
 QY 295 ----- 295
 Db 344 IleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGluTyrIleAla 363
 QY 295 ----- 295
 Db 364 ArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAspArgSerGln 383
 QY 295 ----- 295
 Db 384 ThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValIleGluAspVal 403
 QY 295 ----- 295
 Db 404 MetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeuIleIleTyr 423
 QY 295 ----- 295
 Db 424 GluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPheProGlnSer 443
 QY 295 ----- 295
 Db 444 HisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPheArgHisLeu 463
 QY 295 ----- 295
 Db 464 TyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGlyGlyLeuPro 483
 QY 295 ----- 295
 Db 484 AlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSerGlyGluTrp 503
 QY 296 -----GTTGATGAAGTCACAGGCTCGTATAT 322
 Db 504 GluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArgLeuValTyr 523
 QY 323 TTTGAAGGCACCAAGACTCCCTTTAGAGCATACCTGTACGTAGTCAGTACGTAAAT 382
 Db 524 PheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSerTyrValAsn 543
 QY 383 CTGTGAGAGGTGACAAAGGCTGACTGACCGTGGCTACTCACATTTCTGTCATCAGTCAG 442
 Db 544 ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCysIleSerGln 563
 QY 443 CACTGTGACTTCTTTTATAGTAAAGTATAGTAACACAGAGAATCCACACTGTGTCTCCTT 502

Db	564	HisCysaspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeu	583
Qy	503	TACAAGCTATCAAGTCTCTGAAGATGACCCCAACTTGCAGAAACAAAGAAATTTTGGGCCACC	562
Db	584	TyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPheTyrAlaThr	603
Qy	563	ATTTTGGATTACAGCAGGTCTCTCTGACTATCTCTCCAGAAATTTTCTCTTTTGA	622
Db	604	IleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIlePheSerPheGlu	623
Qy	623	AGTACTACTGGATTACATTGATGGATGCTCTCAAGCCCTCATGATCTACAGCTGGA	682
Db	624	SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly	643
Qy	683	AAGAAATATCTACTGTCTGTTTCATATATATGGTGGTCTCTCAGGTGCAGTTGGTGAATAAT	742
Db	644	LysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeuValAsnAsn	663
Qy	743	CGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCTAGCCCTCTCTAGGTTATGTGGTT	802
Db	664	ArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrValVal	683
Qy	803	GTACTGATAGACACAGGGATCTCTCACCGAGGCTTAAATTTGAAGCGCTTTAAA	862
Db	684	ValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLys	703
Qy	863	TATAAATGGGTCAAATAGAAATTTGACGATCAGGTGGAGGACTCCAATATCTAGCTTCT	922
Db	704	TyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyrLeuAlaSer	723
Qy	923	CGATATGATTTCAATGATCTAGATCTGTGGGATCCACGGCTGCTCTATGAGGATAC	982
Db	724	ArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrSerTyrGlyGlyTyr	743
Qy	983	CTCTCCCTGATGGCATTAATGCAGAGTCAAGTATCTTCAGGTTGCTATTGCTGGGGCC	1042
Db	744	LeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAlaGlyAla	763
Qy	1043	CCAGTCACTCTGTGGATCTTCTATGATACAGGATACACG	1081
Db	764	ProValThrLeuTrpIlePheTyrAspThrGlyTyrThr	776
RESULT 4			
ID	ABG61591	standard; protein; 882 AA.	
XX	AC	ABG61591;	
XX	DE	Human DPPIV related serine protease DPPP-1.	
XX	DE	Human DPPIV related serine protease DPPP-1.	
XX	KW	Human; serine protease; dipeptidyl peptidase IV-related protein; DPPP;	
XX	KW	DPPIV; infection; human immunodeficiency virus; HIV-1; HIV-2; pain;	
XX	KW	diabetes; infertility; obesity; anorexia; Parkinson's disease; stroke;	
XX	KW	heart failure; hypertension; urinary retention; osteoporosis; cancer;	
XX	KW	ulcer; allergy; cancer; psychotic disorder; neurological disorder;	
XX	KW	dyskinesia; reproductive disorder; inflammatory disorder;	
XX	KW	metabolic disorder.	
OS	XX	Homo sapiens.	
XX	PN	WO200231134-A2.	
XX	XX	18-APR-2002.	
PD	XX	12-OCT-2001; 2001WO-US031874.	
PF	XX	12-OCT-2000; 2000US-0240117P.	
XX	FR	(FERR) FERRING BV.	
XX	PA		
XX	XX		

PI	Qi S, Akinsanya KO, Riviere PJ, Junien J;		
XX	WPI; 2002-444178/47.		
DR	N-PSDB; ABK83322.		
XX	New dipeptidyl peptidase IV-related proteins and nucleic acids encoding		
PT	the proteins, useful for treating e.g. fungal, bacterial, protozoan and		
PT	viral infections, cancers, allergies, neurological disorders, or pain.		
XX	Claim 17; Fig 1; 113pp; English.		
XX	The present invention relates to the isolation of novel human serine		
CC	proteases referred to as dipeptidyl peptidase IV (DPPIV)-related proteins		
CC	(DPPP). The dipeptidyl peptidase IV-related proteins (DPPP) and nucleic		
CC	acids encoding them are useful for treating infections such as fungal,		
CC	bacterial, protozoan and viral infections, particularly infections caused		
CC	by human immunodeficiency virus (HIV-1 or HIV-2), pain, diabetes,		
CC	precocious puberty, infertility, obesity, anorexia, bulimia, Parkinson's		
CC	disease, acute heart failure, hypotension, hypertension, urinary		
CC	retention, osteoporosis, angina pectoris, stroke, ulcers, asthma,		
CC	allergies, cancers, migraine, vomiting, psychotic and neurological		
CC	disorders (e.g. anxiety, dementia, or schizophrenia), and dyskinesias.		
CC	These may also be used in discovering therapeutic agents for the		
CC	treatment of reproductive, inflammatory and metabolic disorders. ABG61591		
CC	-ABG61612 represent human DPPP proteins		
XX	Sequence 882 AA;		
SQ	Alignment Scores:		
	Pred. No.: 1.75e-204 Length: 882		
	Score: 1836.50 Matches: 360		
	Percent Similarity: 67.5% Conservative: 0		
	Best Local Similarity: 67.5% Mismatches: 0		
	Query Match: 94.8% Indels: 173		
	DB: 5 Gaps: 1		
	US-10-825-632-8 (1-1083) x ABG61591 (1-882)		
Qy	2 GAAGAAGATGCAGATCAGCTGGAGTGGCTACCTTTGTTCTCCAGAGAATTGATAGA	61	
Db	244 GluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluPheAspArg	263	
Qy	62 TATCTGGCTATTGGTGGTGTCCAAAGCTGAAACCACTCCACGCTGGTGAATTTCTT	121	
Db	264 TyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGlyLysLeu	283	
Qy	122 AGAATTCCTATATGAAGAAATGATGATCTGAGTGAATATTTCATCTTACATCCCT	181	
Db	284 ArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleHisValThrSerPro	303	
Qy	182 ATGTTGGAACAACAGAGGCGAGATTTCATTCCTGTTATCTTAAACAGGTACAGCAATCT	241	
Db	304 MetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThrAlaAsnPro	323	
Qy	242 AAAGTCACTTTTAAGATGTGAGAAATAATGATTGATGCTGAAGAGGATCATATA	295	
Db	324 LysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIleIleAspVal	343	
Qy	295	295	
Db	344 IleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGluTyrIleAla	363	
Qy	295	295	
Db	364 ArgAlaGlyTyrThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAspArgSerGln	383	
Qy	295	295	
Db	384 ThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGluAspAspVal	403	
Qy	295	295	
Db	404 MetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeuIleIleTyr	423	

QY	295	-----	295
Db	424	GluGluThrThrAspIleTrpIleAenIleHisAspIlePheHisValPheProGlnSer	443
QY	295	-----	295
Db	444	HisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPheArgHisLeu	463
QY	295	-----	295
Db	464	TyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGlyGlyLeuPro	483
QY	295	-----	295
Db	484	AlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSerGlyGluTrp	503
QY	296	-----	322
Db	504	GluValLeuGlyArgHisGlySerAenIleGlnValAspGluValArgArgLeuValTyr	523
QY	323	TTTCGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTAGTCAGTTCAGTAAAT	382
Db	524	PheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSerTyrValAsn	543
QY	383	CCTCGAGAGTCACAGGCTGACTGACCGTGGCTACTCACATTCCTGCTGCATCAGTCAG	442
Db	544	ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCysIleSerGln	563
QY	443	CACGTGACTCTTTATAGTAAGTATAGTAACCGAGAGATCCACACTGTGTGTCCTT	502
Db	564	HisCysAspPhePheIleSerLysTyrSerAenGlnLysAenProHisCysValSerLeu	583
QY	503	TACAAGCTATCAAGTCTGAAGATGACCCAACTTCGAAACAAAGAAATTTTGGGCCACC	562
Db	584	TyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPheTrpAlaThr	603
QY	563	ATTTTGGATTCAGCAGCTCCTCTCTGACTACTACTCTCCAGAAATTTTCTCTTTGAA	622
Db	604	IleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIlePheSerPheGlu	623
QY	623	AGTACTACTGGATTTACATTGTATGGATGCTCTACAGCCTCATGATCTACAGCCTGGA	682
Db	624	SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly	643
QY	683	AAGAAATATCTACTGTGCTGTTTCATATATGTGTGCTCTCAGGTGCAGTTCGTGCAATAAT	742
Db	644	LysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeuValAsnAsn	663
QY	743	CGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCTAGCTCTCTAGGTTATGTGGTT	802
Db	664	ArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrValVal	683
QY	803	GTAGTGTACACACAGGGATCTGTCCAGGAGGCTTAATTTGAAGGCGCTTTAAA	862
Db	684	ValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyValaPheLys	703
QY	863	TATTAATGGGTCAATAGAAATGACATCGATCGGTGGAAGGACTCCAAATATCTAGCTTCT	922
Db	704	TyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyrLeuAlaSer	723
QY	923	CGATATGATTTCATCTAGTCTGTGGGCATCCACGGCTGCTCTATGGAGGATAC	982
Db	724	ArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyrGlyGlyTyr	743
QY	983	CTCTCCCTGATGGCATTATGACAGAGTTCAGATATCTTCAGGGTTGCTATTGCTGGGGCC	1042
Db	744	LeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAlaGlyAla	763
QY	1043	CCAGTCTACTCTGTGGATCTCTATGATACAGATACACG	1081
Db	764	ProValThrLeuTrpIlePheTyrAspThrGlyTyrThr	776

RESULT 5	
ID	AAG78415
XX	standard; protein; 882 AA.
AC	AAG78415;
XX	
DT	12-APR-2002 (first entry)
XX	
DE	Amino acid sequence of 21953 human prolyl oligopeptidase.
XX	
KW	21953 prolyl oligopeptidase; antibody; proline; endopeptidase; cancer;
KW	cardiovascular disease; autoimmune disease; atopic allergy;
KW	neuronal disorder; vascular disorder; prostate disorder; cytostatic;
KW	antidiabetic; antiarthritic; antiasthmatic; antiinflammatory;
KW	diabetes mellitus; arthritis; multiple sclerosis; asthma;
KW	Grave's disease; neuronal disorder; demyelinating disease.
XX	
OS	Homo sapiens.
XX	
PN	WO200179473-A2.
XX	
PD	25-OCT-2001.
XX	
PF	11-APR-2001; 2001WO-US040483.
XX	
PR	18-APR-2000; 2000US-0197508P.
XX	
PA	(MILL-) MILLENNIUM PHARM INC.
XX	
PI	Meyers RA, Williamson M;
XX	
DR	WPI; 2002-034353/04.
DR	N-PSDB; AAH99934.
XX	
PT	New polypeptides 21953, member of human prolyl oligopeptidase family,
PT	useful as diagnostic targets and therapeutic agents for controlling
PT	cancer, lymphoma and leukemia.
XX	
PS	Claim 1; Page 102-103; 121pp; English.
XX	
CC	This invention relates to an isolated 21953 human prolyl oligopeptidase.
CC	Which is cystostatic, antidiabetic, antiarthritic, neuroprotective,
CC	antithyroid, dermatological, antipsoriatic, antiasthmatic,
CC	ophthalmological, antiinflammatory, nootropic, antianginal, cardiac,
CC	anticonvulsant, gynaecological, vasotrophic, antiParkinsonian,
CC	antiatherosclerotic, anorectic and metabolic in its action. Uses include
CC	gene therapy, expression or activity of 21953 protein modulator, it is
CC	useful for identifying a compound which binds to it and can be used in
CC	preventing, treating or detecting a cellular proliferative or
CC	differentiative disorder. The 21953 molecules can act as novel diagnostic
CC	targets and therapeutic agents for controlling disorders associated with
CC	the aberrant activity or degradation of peptide hormones e.g., disorders
CC	associated with cell differentiation and proliferation such as cancer
CC	immune function, reproductive, neurological and cardiovascular function.
CC	The 21953 molecules are thus useful for treating and preventing cellular
CC	proliferative and differentiative disorders, haematopoietic neoplastic
CC	disorders, immune disorders such as autoimmune diseases, diabetes
CC	mellitus, arthritis, multiple sclerosis, asthma, Grave's disease,
CC	neuronal disorders, demyelinating diseases, vascular disorders and
CC	metabolism or pain disorders. This sequence represents the amino acid
CC	sequence of 21953 human prolyl oligopeptidase
XX	
SQ	Sequence 882 AA;

Alignment Scores:			
Pred. No.:	1.75e-204	Length:	882
Score:	1836.50	Matches:	360
Percent Similarity:	67.5%	Conservative:	0
Best Local Similarity:	67.5%	Mismatches:	0
Query Match:	94.8%	Indels:	173
DB:	5	Gaps:	1
US-10-825-632-8 (1-1083) x AAG78415 (1-882)			

QY	2	GAGAAGATCCAGATCAGCTGAGTCGCTACTTGTCTCCCAAGAAATTTGATAGA	61
Db	244	GlulGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluPheAspArg	263
QY	62	TAATCTGGCTATTGGTGGTCTCAAAAGCTGAACAACCTCCAGTGGTGGTAAATCTT	121
Db	264	TyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGlyValLeu	283
QY	122	AGAATTCATATGAAGAAATGATGAATCTGAGGTGGAAATTAATTCATGTCCTCC	181
Db	284	ArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleHisValThrSerPro	303
QY	182	ATGTGGAAACAAGAGGGAGATTCATCCGTATCTAAACAGGTACAGCAATCCT	241
Db	304	MetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThrAlaAsnPro	323
QY	242	AAAGTCATTTTAAGATGTGAGAAATAATGATTGATGCTGAAGGAAGGATCATA	295
Db	324	LysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIleIleAspVal	343
QY	295	-----	295
Db	344	IleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGluTyrIleAla	363
QY	295	-----	295
Db	364	ArgAlaGlyTyrThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAspArgSerGln	383
QY	295	-----	295
Db	384	ThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGluAspAspVal	403
QY	295	-----	295
Db	404	MetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeuIleIleTyr	423
QY	295	-----	295
Db	424	GluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPheProGlnSer	443
QY	295	-----	295
Db	444	HisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPheArgHisLeu	463
QY	295	-----	295
Db	464	TyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGlyLeuPro	483
QY	295	-----	295
Db	484	AlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSerGlyGluTrp	503
QY	296	-----GTTGATGAAGTCAGAACGCTGGTATAT	322
Db	504	GluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgLeuValTyr	523
QY	323	TTTGAAGGCACCAAGATCCCTTTAGAGCATCACCTGTAGTACGTAGTCAGTAAAT	382
Db	524	PheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSerTyrValAsn	543
QY	383	CCTGGAGGTGACAAGGTGACTGACCGTGGCTACTCACATTCCTGTCATCAGTCAG	442
Db	544	ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysIleSerGln	563
QY	443	CACGTGACTCTTTTATAAGTAAAGTAAACCAAGAAATCCACACTGTGTGCTCCTT	502
Db	564	HisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeu	583
QY	503	TACAGTATCAGTCTCTGAGATGACCAACTGCAAAACAAGAAATTTGGGCCACC	562
Db	584	TyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPheTrpAlaThr	603
QY	563	ATTTTGGATTACAGAGGTCTCTTCTGACTATACTCTCCAGAAATTTTCTTTGAA	622
Db	604	IleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIlePheSerPheGlu	623
QY	623	AGTACTACTGGATTACATTGATGGATGCTCTCAAGCCTCATGATCTACAGCCTCGA	682
Db	624	SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly	643
QY	683	AAGAAATATCTACTGCTGCTTCAATATATGTTGGTGGTCTCAGGTGCAGTGTGTAATAT	742
Db	644	LysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeuValAsnAsn	663
QY	743	CGTTTAAAGAGTCAAGATTTTCCGCTTCAATACCTAGCTCTCTCTAGTTTATGTGTT	802
Db	664	ArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrValVal	683
QY	803	CTAGTGTAGTACACACAGGGATCTCTGCACCGAGGCTTAAATTTGAAGCGCCTTTAA	862
Db	684	ValValIleAspAsnAspGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLys	703
QY	863	TATAAAATGGGTCAAATAGAAATTCAGCATCAGGTGGAAGGATCCCAATATCTAGCTTCT	922
Db	704	TyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyrLeuAlaSer	723
QY	923	CGATATGATTTTCATTGACTTAGATCGTGGGCATCCACGCTGGTCCCTATGAGGATAC	982
Db	724	ArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyrGlyGlyTyr	743
QY	983	CTCTCCCTGATGGCATTAAATGCAGAGTTCAGATATCTTCAGGGTTCCTATTGCTGGGCC	1042
Db	744	LeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAlaGlyAla	763
QY	1043	CCAGTCACCTGTGTGATCTTCTATGATACAGATACAGCAGCAGCAGCAGCAGCAGC	1081
Db	764	ProValThrLeuTrpIlePheTyrAspThrGlyTyrThr	776
RESULT 6			
ID	AAE24170 standard; protein; 882 AA.		
AC	AAE24170;		
DT	23-SEP-2002 (first entry)		
DE	Human dipeptidyl peptidase 8 (DPP8) protein.		
KW	Human; dipeptidyl peptidase; DPP; neoplasia; type II diabetes; cirrhosis; autoimmunity; human immuno deficiency virus; HIV infection; cycostatic; graft rejection; antidiabetic; antiinflammatory; immunosuppressive; antiviral; enzyme.		
OS	Homo sapiens.		
PN	WO200234900-A1.		
PD	02-MAY-2002.		
PF	29-OCT-2001; 2001WO-AU001388.		
PR	27-OCT-2000; 2000AU-00001078.		
PA	(UNSW) UNIV SYDNEY.		
PI	Abbott CA, Gorrell MD;		
DR	WPI; 2002-454646/48.		
XX	N-PSDB; AAD38956.		
PT	New dipeptidyl peptidase (Dpp) peptides, useful for screening inhibitors of DPP catalytic activity, which may be employed to treat e.g. neoplasia, type II diabetes, cirrhosis, autoimmunity, graft rejection and HIV infection.		

PS Example; Fig 1; 91pp; English.

XX The present invention relates to dipeptidyl peptidase (DPP) proteins and polynucleotides encoding such proteins. The DPP peptides are useful for screening inhibitors of DPP catalytic activity. The inhibitors are useful for treating neoplasia, type II diabetes, cirrhosis, autoimmunity, graft rejection and HIV (human immunodeficiency virus) infection. The present sequence is human DPP8 protein

XX SQ Sequence 882 AA;

Alignment Scores:
Pred. No.: 1-75e-204 Length: 882
Score: 1836.50 Matches: 360
Percent Similarity: 67.5% Conservative: 0
Best Local Similarity: 67.5% Mismatches: 0
Query Match: 94.8% Indels: 173
DB: 5 Gaps: 1

US-10-825-632-8 (1-1083) x AAE24170 (1-882)

QY 2 GRAGAGATGCCAGATCAGCTGGAGTCCTACCTTTTCTTCCCAAGAAGATTGTAGA 61
DB 244 GIUGLUSPALARGSERALAGLYVALALATHRPHEVALLEUGLUGLUPHEASPARG 263
QY 62 TATTCGGCTATGTGTGTGCCAAAGCTGAAACACTCCAGTGGTGGTAAATTCCT 121
DB 264 TYRSERGLYTYRTRTPCYSPROLYSALAGLUTHRTHRPROSERGLYGLYLYSILELEU 283
QY 122 AGAATCTATATGAAGAAATGATGAATCTGAGTGGAAATTTATCATGTTACATCCCT 181
DB 284 ARGILELEUTYRGLUGLUSNAPGLUSERGLUVALGLUILEILEHISVALTHRSEPRO 303
QY 182 ARGTTGGAACAAGAGGAGGAGATTCATTCCTGTTATCCTAAACAGTACACCAATCCT 241
DB 304 METLEUGLUTHRARGALASPSPHEARGTYRPROLYSTHRGLYTHRALASNP 323
QY 242 AAGTCACTTTAAGATGTCAGAAATATGATGCTGAGGAGGATCATATA 295
DB 324 LYSVALTHRPHELYSMETSERGLUILEMETILEASPALAGLUGLYARGILEILEASPV 343
QY 295 ----- 295
DB 344 ILEASPLYSGLULEULEGINPROPHLEGLULELEUPHEGLUVALGLUTYRILEALA 363
QY 295 ----- 295
DB 364 ARGALAGLYTRTPRPROGLUGLYLYSTYRVALATPSERILELEULEASPARSERGLN 383
QY 295 ----- 295
DB 384 THRARGLEUGLINILEVALLEUSERPROGLULEUPHEILEPROVALGLUASPV 403
QY 295 ----- 295
DB 404 METGLUARGGLNARGLEULEGLUSERVALPROASPSERVALTHRPROLEULEITYR 423
QY 295 ----- 295
DB 424 GLUGLUTHRTHRASPLETERPLEASNILEHISASPILEPHEHISVALPHEPROGLN 443
QY 295 ----- 295
DB 444 HISGLUGLUGLULEGLUPHELEPHEALASERGLUCYSLYSTHRGLYPHEARGHISLEU 463
QY 295 ----- 295
DB 464 TYRILELETHRSETRILELEULYSGLUSERTYRLYSARGSERGLYGLYLEUPRO 483
QY 295 ----- 295
DB 484 ALAPROSERASPHELYSCYSPROILELYSGLULEALALLETHRSETRILEGLUTP 503

QY 296 -----? GTTGATGAAGTCAGAGGCTGCTATAT 322
DB 504 GLUVALLEUGLYARGHISGLYSERANILIEGLINVALASPGLUVALARGLEUVALTYR 523
QY 323 TTGAAGGACCAAGACATCCCTTTAGAGATCATACCTGTAGTACGTAGTACGTAAAT 382
DB 524 PHEGLUGLYTHRLYSASPSPROLEUGLUIHISLEUTYRVALVALSERTYRVALASN 543
QY 383 CTGGAGAGGTCAAGGCTGACTGACCGTGGCTACTCACATTTCTGCTGCATCAGTCAG 442
DB 544 PROGLYGLUVALTHRARGLEUTHRASPARGGLYTYRSERHISSERCYSEILESERGLN 563
QY 443 CACTGTGACTTCTTTATAAGTAAAGTATAGTAACACAGAGAAATCCACACTGTGCTCCCT 502
DB 564 HISCYASPPHEPHEILESERLYSTYRSERANGLINLYSASNPROHISCYSEVALSERLEU 583
QY 503 TACAAGCTTCAAGTCCCTGAGATGACCAACTTGCACAAAGAAAGAAATTTGGGCCACC 562
DB 584 TYRLEULEUSERSERPROGLUASPPSPROTHRCYSLYSTRHLYSGLUPHETTPALATHR 603
QY 563 ATTTTGGATTACAGAGGTCCTCTTCTGACTATATCTCTCCAGAAATTTCTCTTTTGA 622
DB 604 ILELEUASPSERALAGLYPROLEUPROASPLYRTHRPROGLUILEPHESERPHEGLU 623
QY 623 AGTACTACTCGGATTTACATTTGATGGATGCTCTACAGCCTCATGATCTACAGCCTGGA 682
DB 624 SERTHRGLYPHETHRLEUTYRGLYMETLEUTYRLYSPROHISASPLEUGLINPROGLY 643
QY 683 AAGAAATATCTACTGTCTCTTTCATATATGTTGGTCTCTCAGTGCGAGTTGGTGAATAAT 742
DB 644 LYSLYSTYRPROTHRVALLEUPHEILETYRGLYGLYPROGLINLEUVALASN 663
QY 743 CGGTTTAAAGGAGTCAAGTATTTCCGTTGAATACCTACCTAGCCTCTCTAGGTTATGGTT 802
DB 664 ARGPHELYSGLYVALLYSTYRPHARGLEUASNTHRLEUASERLEUGLYTYRVALVAL 683
QY 803 GTAGTGATAGACAAACAGGGATCTGTCCAGGAGGCTTAAATTTCAAGGCGCTTTAAA 862
DB 684 VALVALILEASPAENARGGLYSERCYHISARGGLYLEULYSPHEGLUGLYALAPHELYS 703
QY 863 TATAAATGGTCAAAATAGAAATTCAGATCAGGTGGAAGGACTCCCAATATCTAGTCTCT 922
DB 704 TYRLYSMETGLYGLINILEGLUILEASPSGLNVALGLUGLYLEUGLINTYRLEUASER 723
QY 923 CGATATGATTTTCATTTAGCTTAGATCGTGTGGGATCCAGGCTGGTCTCTATCGAGATAC 982
DB 724 ARGTYRASPHEILEASPLEUASPARGVALGLYILEHISGLYTRPSERTYRGLYTYR 743
QY 983 CTCTCCCTGATGGCATTAAATGCAGAGTACAGATATCTTCAGGTTGCTATTCTGGGCC 1042
DB 744 LEUSERLEUMETALALEUMMETGLNARGSERASPILEPHEARGVALAILEALAGLYALA 763
QY 1043 CCAGTCACTCTGTGGATCTTCTATGATACAGGATACACG 1081
DB 764 PROVALTHRLEUTRIPLEPHETYRASPTHRGLYTYRTHR 776
RESULT 7
AAU74749
ID AAU74749 standard; protein; 882 AA.
XX
AC AAU74749;
XX
DT 09-APR-2002 (first entry)
XX
DE Human protease PRTS-9 protein sequence.
XX
KW Human; protease; PRTS; gastrointestinal; Crohn's disease; cancer;
KW cardiovascular; atherosclerosis; autoimmune disorder; dermatitis;
KW inflammatory disorder; acquired immunodeficiency syndrome; AIDS;
KW cell proliferative disorder; developmental disorder; epilepsy;
KW Duchenne muscular dystrophy; epithelial disorder; neurological disorder;
KW reproductive disorder; endometriosis.

OS Homo sapiens.
 PN WO200198468-A2.
 XX
 PD 27-DEC-2001.
 XX
 PF 13-JUN-2001; 2001WO-US019178.
 XX
 PR 16-JUN-2000; 2000US-0212336P.
 PR 22-JUN-2000; 2000US-0213955P.
 PR 29-JUN-2000; 2000US-0215396P.
 PR 07-JUL-2000; 2000US-0216821P.
 PR 14-JUL-2000; 2000US-0218946P.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 XX Yue H, Elliott VS, Gandhi AR, Lal P, Au-Young J, Tribouley CM,
 PI Deleage AM, Baughn MR, Nguyen DB, Lee EA, Hafalia A, Khan FA;
 PI Walia NK, Yao MG, Lu DAM, Patterson C, Tang YT, Walsh RT;
 PI Azimzai Y, Lu Y, Ramkumar J, Xu Y, Reddy R, Das D, Kearney L;
 PI Kallick DA;
 XX
 DR WPI: 2002-090437/12.
 DR N-PSDB; ABK12892.
 XX
 XX Twenty one human proteases (referred to as PRTS-1 to PRTS-21), useful in
 PT the diagnosis, treatment and prevention of gastrointestinal (e.g.
 PT gastritis), cardiovascular (e.g. atherosclerosis) and cell proliferative
 PT (e.g. cancer) disorders.
 XX
 PS Claim 1; Page 140-142; 177pp; English.
 XX
 CC The present invention relates to twenty one new human proteases, referred
 CC to as PRTS-1 to PRTS-21. The PRTS polynucleotides and polypeptides of the
 CC invention are useful in the diagnosis, treatment and prevention of
 CC gastrointestinal e.g. gastritis, esophageal carcinoma and Crohn's
 CC disease, cardiovascular e.g. atherosclerosis, hypertension and myocardial
 CC infarction, autoimmune/inflammatory e.g. acquired immunodeficiency
 CC syndrome (AIDS), allergies and osteoarthritis, cell proliferative e.g.
 CC cancer, developmental e.g. Duchenne and Becker muscular dystrophy,
 CC epithelial e.g. dermatitis, neurological e.g. epilepsy and Alzheimer's
 CC disease and reproductive e.g. infertility and endometriosis disorders.
 CC Numerous other examples of each disorder are given in the specification.
 CC The present protein sequence represents the human protease PRTS-9 protein
 CC of the invention
 XX
 SQ Sequence 882 AA;
 Alignment Scores:
 Pred. No.: 1.75e-204 Length: 882
 Score: 1836.50 Matches: 360
 Percent Similarity: 67.5% Conservative: 0
 Best Local Similarity: 67.5% Mismatches: 0
 Query Match: 94.8% Indels: 173
 DB: 5 Gaps: 1
 US-10-825-632-8 (1-1083) x AAU74749 (1-882)
 QY 2 GAAGAATGCCAGATCAGCTGGAGTCGCTACCTTGTCTCCAAAGAAGATTGTAGA 61
 Db 244 GluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGluGluPheAspArg 263
 QY 62 TATTCTGGCTATTGGTGGTCCAAAGCTGAAACAACTCCCGAGTGGTGAATTCCT 121
 Db 264 TyrSerGlyTyrTrpCysProLysAlaGluThrThrProSerGlyGlyLysLeu 283
 QY 122 AGAATTCATATGAAGAAATGATGATCTCGAGGTGGAATATTATCATGTTACATCCCT 181
 Db 284 ArgLeuLeuTyrGluGluAsnAspGluSerGluValGluLeuLeuHisValThrSerPro 303
 QY 182 ATGTGTGAACAAGAGGGCAGATTTCATTCCTCGTTATTCCTAAACAGGTACAGCAATCCT 241
 Db 304 MetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThrAlaAsnPro 323

QY 242 AAAGTCACCTTTAAGATGTCAGAAATAATGATGATGCTGAAGGAAGATCATA----- 295
 |||||
 Db 324 LysValThrPheLysMetSerGluLeuMetIleAspAlaGluGlyArgIleAspVal 343
 |||||
 QY 295 ----- 295
 Db 344 IleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGluTyrIleAla 363
 |||||
 QY 295 ----- 295
 Db 364 ArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAspArgSerGln 383
 |||||
 QY 295 ----- 295
 Db 384 ThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGluAspAspVal 403
 |||||
 QY 295 ----- 295
 Db 404 MetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeuIleIleTyr 423
 |||||
 QY 295 ----- 295
 Db 424 GluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPheProGlnSer 443
 |||||
 QY 295 ----- 295
 Db 444 HisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPheArgHisLeu 463
 |||||
 QY 295 ----- 295
 Db 464 TyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGlyGlyLeuPro 483
 |||||
 QY 295 ----- 295
 Db 484 AlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSerGlyGluTrp 503
 |||||
 QY 296 -----GTTGTAAGTACAGAGCGTGTATAT 322
 |||||
 Db 504 GluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgLeuValTyr 523
 |||||
 QY 323 TTTGAAGGCACCAAGACCTCCCTTTAGAGCATCACCTGTACGTAGTCAGTTACGTAAT 382
 |||||
 Db 524 PheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSerTyrValAsn 543
 |||||
 QY 383 CCTCGAGAGTGACAAAGGCTGACTGACCGTGGCTACTCATCTTCTGTCGATCAGTCAG 442
 |||||
 Db 544 ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCysIleSerGln 563
 |||||
 QY 443 CACTGTGACTTCTTTATAGTATAGTATACCAAGAGAAATCCACACTGTGTCCCTT 502
 |||||
 Db 564 HisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeu 583
 |||||
 QY 503 TACAAGCTATCAAGTCTCTGAAGATGACCCAACTTGCAAAACAAAGAAATTTGGGCCACC 562
 |||||
 Db 584 TyrLysLeuSerSerProGluAspAspPProThrCysLysThrLysGluPheTrpAlaThr 603
 |||||
 QY 563 ATTTTGGATTACAGAGGTCCTCTTCTGACTATATCTCTCCAGAAATTTCTCTTTTGA 622
 |||||
 Db 604 IleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePheSerPheGlu 623
 |||||
 QY 623 AGTACTACTGGATTACATTGTATGGGATGCTCTACAGCCTCATCATCTACAGCCTCGA 682
 |||||
 Db 624 SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly 643
 |||||
 QY 683 AAGAAATATCTACTGTGCTGTTTCATATATATGTTGGTCTCTCAGTGCAGTGTGGTGAATAT 742
 |||||
 Db 644 LysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeuValAsnAsn 663
 |||||
 QY 743 CGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCTAGCTCTCTAGTTATCTGTT 802
 |||||
 Db 664 ArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrValVal 683

QY 803 GTAGTATAGACAACAGGGGATCCTCTCACCAGGCGCTTAAATTTGAAGGCGCTTTAAA 862
Db ValValIleAaspAenArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLys 703
QY 863 TATAAATGGTCAATAGAAATTGACGATCAGGTGGAAGGACCTCCATATCTAGCTTCT 922
Db TyrlsMetGlyGlnIleGluIleAaspGlnValGluGlyLeuGlnTyrlsLeuLaser 723
QY 923 CGATATGATTTTCATTCAGCTAGATCGTGTGGGCATCCAGGCTGGTCTATGGAGGATAC 982
Db ArgTyrAaspPheIleAaspLeuAaspGlnValGlyIleHisGlyTyrPserTyrGlyGlyTyr 743
QY 983 CTCTCCCTGATGCATTATGACGAGGTGAGATATCTTCAGGGTGTCTATTGCTGGGGCC 1042
Db LeuSerLeuMetAlaLeuMetGlnArgSerAaspIlePheArgValAlaIleAlaGlyAla 763
QY 1043 CCACTACTCTGTGGATCTTCTATGATACAGGATACACG 1081
Db ProValThrLeuTrpIlePheTyrAaspThrGlyTyrThr 776
RESULT 8
ADII17086
ID ADII17086 standard; protein: 882 AA.
XX AC ADII17086;
XX DT
XX DE 15-APR-2004 (first entry)
XX DE Human NOVX protein homologue SeqID 622.
XX KW human; NOVX; cardiomyopathy; atherosclerosis; cancer; diabetes;
XX KW inflammation; autoimmune disorder; allergy; blood disorder;
XX KW acquired immunodeficiency syndrome; AIDS; obesity; asthma;
XX KW immunoglobulin (Ig)A nephropathy; cirrhosis; arthritis;
XX KW Alzheimer's disease; infection; str.
XX OS Homo sapiens.
XX FN WO200268649-A2.
XX PD 06-SEP-2002.
XX PF 31-JAN-2002; 2002WO-US0003785.
XX PR 31-JAN-2001; 2001US-0265395P.
XX PR 31-JAN-2001; 2001US-0265412P.
XX PR 31-JAN-2001; 2001US-0265514P.
XX PR 31-JAN-2001; 2001US-0265517P.
XX PR 02-FEB-2001; 2001US-0266406P.
XX PR 05-FEB-2001; 2001US-0266767P.
XX PR 07-FEB-2001; 2001US-0266975P.
XX PR 07-FEB-2001; 2001US-0267057P.
XX PR 08-FEB-2001; 2001US-0267459P.
XX PR 09-FEB-2001; 2001US-0267823P.
XX PR 15-FEB-2001; 2001US-0268974P.
XX PR 26-FEB-2001; 2001US-0271564P.
XX PR 27-FEB-2001; 2001US-0271853P.
XX PR 27-FEB-2001; 2001US-0271955P.
XX PR 02-MAR-2001; 2001US-0273046P.
XX PR 02-MAR-2001; 2001US-0273046P.
XX PR 14-MAR-2001; 2001US-0275925P.
XX PR 14-MAR-2001; 2001US-0275947P.
XX PR 14-MAR-2001; 2001US-0275950P.
XX PR 14-MAR-2001; 2001US-0275989P.
XX PR 15-MAR-2001; 2001US-0276448P.
XX PR 16-MAR-2001; 2001US-0276450P.
XX PR 16-MAR-2001; 2001US-0276397P.
XX PR 20-MAR-2001; 2001US-0276768P.
XX PR 20-MAR-2001; 2001US-0276652P.
XX PR 26-MAR-2001; 2001US-0278775P.
XX PR 26-MAR-2001; 2001US-0278775P.
XX PR 29-MAR-2001; 2001US-0279882P.
XX PR 29-MAR-2001; 2001US-0279884P.

PR 30-MAR-2001; 2001US-0280147P.
PR 11-APR-2001; 2001US-0282992P.
PR 11-APR-2001; 2001US-0383083P.
PR 20-APR-2001; 2001US-0385133P.
PR 23-APR-2001; 2001US-0385749P.
PR 03-MAY-2001; 2001US-0288327P.
PR 03-MAY-2001; 2001US-0288504P.
PR 29-MAY-2001; 2001US-0294047P.
PR 30-MAY-2001; 2001US-0294473P.
PR 08-JUN-2001; 2001US-0296964P.
PR 18-JUN-2001; 2001US-0298959P.
PR 13-AUG-2001; 2001US-0299324P.
PR 16-AUG-2001; 2001US-0312020P.
PR 16-AUG-2001; 2001US-0312889P.
PR 21-AUG-2001; 2001US-0312908P.
PR 21-AUG-2001; 2001US-0313390P.
PR 28-AUG-2001; 2001US-0315470P.
PR 31-AUG-2001; 2001US-0316447P.
PR 07-SEP-2001; 2001US-0318115P.
PR 12-SEP-2001; 2001US-0318118P.
PR 19-SEP-2001; 2001US-0318740P.
PR 18-OCT-2001; 2001US-0320245P.
PR 18-OCT-2001; 2001US-0330308P.
PR 14-NOV-2001; 2001US-0332701P.
XX (CURA-) CURAGEN CORP.
XX Tchernev VT, Spyttek KA, Zerhuseen BD, Patturajan M, Shimkets RA;
XX Li L, Gangolli EA, Padigar M, Anderson DM, Rastelli L, Miller CE;
XX Gerlach VL, Taupier RJ, Gusev VV, Colman SD, Wolenc AR, Pena CEA;
XX Furtak K, Grosse WM, Alsobrook JP, Lepley DM, Rieger DK, Burgess CE;
XX WPI; 2002-706998/76.
XX New NOVX polypeptides and nucleic acids, useful for preventing or
XX treating NOVX-associated disorders, e.g. cancer, cardiomyopathy,
XX atherosclerosis, or diabetes, and in chromosome mapping, tissue typing or
XX pharmacogenomics.
XX Disclosure; SEQ ID NO 622; 1498pp; English.
XX This invention relates to a novel nucleic acids, and encoded polypeptides
XX thereof, which have properties related to the stimulation of biochemical
XX or physiological responses in a cell, tissue, organ or organism.
XX Specifically, it refers to the use of biologically active fragments for
XX diagnostic and prognostic assays and furthermore in the treatment of
XX diverse pathological conditions. The present invention describes novel
XX human and murine NOVX proteins, as well as methods to modulate their
XX expression using antisense oligos, ribozymes and peptide nucleic acids.
XX The NOVX polypeptides, polynucleotides and antibodies are useful in
XX treating or preventing NOVX-associated disorders, e.g. cardiomyopathy,
XX atherosclerosis, cancer and diabetes. Furthermore, they may be used in
XX treating or preventing diseases such as inflammation, autoimmune
XX disorders, allergies, asthma, immunoglobulin (Ig)A nephropathy, cirrhosis
XX (AIDS), obesity, blood disorders, acquired immunodeficiency syndrome
XX arthritis, Alzheimer's disease, infections, stroke, muscular dystrophy
XX and epilepsy, accordingly, these molecules have many activities including
XX cytoskeletal, cardiac and inflammatory, immunosuppressive, antiallergic,
XX haemostatic, anti-HIV, antidiabetic, antiarteriosclerotic, anorectic,
XX anasthmatic, nephrotropic, antirheumatic, hepatotropic,
XX neuroprotective, nootropic, antibacterial, virucide, antiparasitic,
XX relaxant and anticonvulsant. In addition, they are useful in screening
XX assays to identify small molecules that modulate or inhibit, for example,
XX neurogenesis, wound healing and angiogenesis. The nucleic acids are also
XX used as in chromosome mapping, tissue typing, preventive medicine and
XX pharmacogenomics. This polypeptide is a homologue of a human NOVX protein
XX of the invention.
SQ Sequence 882 AA;
Alignment Scores:
Pred. No.: 1.75e-204 Length: 882

Score: 1836.50 Matches: 360
Percent Similarity: 67.5% Conservative: 0
Best Local Similarity: 67.5% Mismatches: 0
Query Match: 94.8% Indels: 173
DB: 5 Gaps: 1

US-10-825-632-8 (1-1083) x AD117086 (1-882)

```
QY 2 GAAGAAGATGCACATCAGCTGGAGTCGCTACCTTTGTTCTCCAAGAAGAAATTTGATAGA 61
DB 244 GluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluPheAspArg 263
QY 62 TATTCTGGCTATTGCTGCTCCAAAAGCTGAAAACAACTCCACAGTGGTGGTAAATTCCTT 121
DB 264 TyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGlyValLeu 283
QY 122 AGAATCTATATGAAGAAATGATGAATCTGAGGTGGAATTAATTCATGTTACATCCCT 181
DB 284 ArgileLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisValThrSerPro 303
QY 182 ATGTTGGAACACAGAGGGCAGATTTCATTCGCTTATCTTAAACAGGTACAGCAATCCT 241
DB 304 MetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThrAlaAsnPro 323
QY 242 AAAGTCACTTTAAGATGTCAGAAATAATGATTGATGCTGAAGGAGGATCATATA 295
DB 324 LysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIleIleAspVal 343
QY 295 ----- 295
DB 344 IleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGluTyrIleAla 363
QY 295 ----- 295
DB 364 ArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAspArgSerGln 383
QY 295 ----- 295
DB 384 ThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGluAspVal 403
QY 295 ----- 295
DB 404 MetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeuIleIleTyr 423
QY 295 ----- 295
DB 424 GluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPheProGlnSer 443
QY 295 ----- 295
DB 444 HisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPheArgHisLeu 463
QY 295 ----- 295
DB 464 TyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGlyGlyLeuPro 483
QY 295 ----- 295
DB 484 AlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSerGlyGluTrp 503
QY 296 -----GTTGATGAAGTCAGAGCTCGGTATAT 322
DB 504 GluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArgLeuValTyr 523
QY 323 TTTGAAGGACCAAGACTCCCTTTAGAGCATACCTGATAGTACGTAGTACGTAGTAAAT 382
DB 524 PheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSerTyrValAsn 543
QY 383 CCTGGAGGTGACAGGCTGACTGACCGTGGCTACTCACATTCCTGCTGTCAGTCAG 442
DB 544 ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCysIleSerGln 563
QY 443 CACTGTGACTTCTTTAAGTAAGTATAGTAACCAAGAAATCCACACTGTGTGCTT 502
```

```
DB 564 HisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeu 583
QY 503 TACAAGCTATCAAGTCTGAGTGCACCCAACTTGCACAAACAAAGAAATTTTGGCCACC 562
DB 584 TyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPheTrpAlaThr 603
QY 563 ATTTTGGATTTCAGCAGGTCCTCTTCCCTGACTACTACTCTCCAGAAATTTTCTTTTGA 622
DB 604 IleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePheSerPheGlu 623
QY 623 AGTACTACTGGATTACATTTGATGGATGCTCTCAAGCCTCATGATCTACAGCCTGGA 682
DB 624 SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly 643
QY 683 AAGAAATATCTACTGCTGCTTCATATATATGTTGGTCTCAGGTGAGTGTGTGTAATAT 742
DB 644 LysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeuValAsnAsn 663
QY 743 CGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCTAGCCTCTCTAGTTATGTGTT 802
DB 664 ArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrValVal 683
QY 803 GTAGTGTAGTACAAACAGGGGATCTCTGCACCGAGGCTTAAATTTGAAGCGCCTTTAAA 862
DB 684 ValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLys 703
QY 863 TATAAATGGTCAAAATAGAAATTTGACGATCAGTGGAGGACTCCAAATATCTAGCTTCT 922
DB 704 TyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyrLeuAlaSer 723
QY 923 CGATATGATTTTCATTTAGATCTGTGGGCATCCAGCGCTGCTTATCGAGGATAC 982
DB 724 ArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyrGlyTyr 743
QY 983 CTCCTCCCTGATGGCATTAAATGCAGAGTCAGATATCTTCAGGTGTCTATTCCTGGGCC 1042
DB 744 LeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAlaGlyAla 763
QY 1043 CCAGTCACTCTGTGGATCTTCTATGATACAGGATACAG 1081
DB 764 ProValThrLeuTrpIlePheTyrAspThrGlyTyrThr 776
RESULT 9
ABU07720
ID ABU07720 standard; protein; 882 AA.
XX
AC ABU07720;
XX
DT 19-MAY-2003 (first entry)
XX
DE Human serine protease HIPHUM46.
XX
KW Human; enzyme; HIPHUM46; serine protease; gene therapy; osteoarthritis;
KW serine protease activity modulation; dipeptidyl peptidase activity;
KW musculoskeletal disease; Hepatitis B virus infection; myotonic dystrophy;
KW Alzheimer's disease; paraneuronal palsy; Huntington's disease;
KW amyotrophic lateral sclerosis; malabsorption syndrome; lung disease;
KW irritable bowel syndrome; type I diabetes; faecal incontinence; anaemia;
KW haemorrhoid; proctitis; rectal polyp; small bowel tumour; dyslexia;
KW colorectal tumour; ceroid lipofuscinosis; allergic encephalomyelitis;
KW multiple sclerosis.
XX
OS Homo sapiens.
XX
Key Location/Qualifiers
FH Region 259..260
FT /note= "Paired glutamates of the beta propeller domain"
FT Active-site 739
FT /label= Catalytic_serine_residue
FT Active-site 817
FT /label= Catalytic_aspartate_residue
FT Active-site 849
```

/label= Catalytic histidine residue

FT		/label= Catalytic_histidine_residue	
XX	GB2374869-A.		
PN			
XX	30-OCT-2002.		
PD			
XX	22-JAN-2002; 2002GB-00001404.		
PF			
XX	23-JAN-2001; 2001GB-00001760.		
PR	(GLAX) GLAXO GROUP LTD.		
XX			
PA			
XX			
PI	Edbrooke MR, Lewis AP;		
XX			
DR	WPI; 2003-150703/15.		
DR	N-PSDB; ABX12255.		
XX			
PT	Identifying modulators of serine protease activity useful for treating		
PT	musculoskeletal diseases, by contacting cell expressing a novel serine		
PT	protease polypeptide with a compound and monitoring serine protease		
PT	activity.		
XX			
PS	Claim 10; Page 26-29; 38pp; English.		
XX			
CC	The invention relates to a method of identifying a substance that		
CC	modulates serine protease activity, comprising contacting a cell such as		
CC	a neuronal cell, lung cell, intestinal cell or a cell infected with a		
CC	virus, expressing a serine protease polypeptide (HIPHUM 46), or its		
CC	variant having dipeptidyl peptidase activity, or a serine protease		
CC	isolated from the cell with a test substance and monitoring for serine		
CC	protease activity. The method is useful for identifying a substance that		
CC	modulates serine protease activity. A modulator of the serine protease is		
CC	useful in the manufacture of a medicament for treatment or prophylaxis of		
CC	a musculoskeletal disease e.g. osteoarthritis, Hepatitis B virus		
CC	infection, Alzheimer's disease, paraneuronal palsy, myotonic		
CC	dystrophy, Huntington's disease or amyotrophic lateral sclerosis.		
CC	Additional disease that may be treated using modulators of the serine		
CC	protease include malabsorption syndromes, irritable bowel syndrome, lung		
CC	disease, type I diabetes, faecal incontinence, haemorrhoids, proctitis,		
CC	rectal polyps, small bowel tumours, colorectal tumours, anaemia,		
CC	dyslexia, ceroid lipofuscinosis, allergic encephalomyelitis, and multiple		
CC	sclerosis. The present sequence represents the amino acid sequence of the		
CC	human serine protease HIPHUM46		
XX			
SQ	Sequence 882 AA;		
	Alignment Scores:		
	Pred. No.:	1,75e-204	Length: 882
	Score:	1836.50	Matches: 360
	Percent Similarity:	67.5%	Conservative: 0
	Best Local Similarity:	67.5%	Mismatches: 0
	Query Match:	94.8%	Indels: 173
	DB:	6	Gaps: 1
	US-10-825-632-8 (1-1083) x ABU07720 (1-882)		
QY	2 GAAGAAGATGCCAGATCAGCTGGAGTCGGCTACCTTTGTTCTCCAAAGAGAATTTGATAGA	61	
Db	244 GluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGluPheAspArg	263	
QY	62 TATTCTGGCTATTGGTGGTCTCCAAAGCTGAAACAACTCCCAAGTCGGTGGTAAATTTCTT	121	
Db	264 TyrSerGlyTyrTrpTyrCysProLysAlaGluThrThrProSerGlyGlyLysIleLeu	283	
QY	122 AGAATTCATTAAGAAGAAATGATGAATCTGAGGTGGAATATTTCATGTTTACATCCCT	181	
Db	284 ArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisValThrSerPro	303	
QY	182 ATGTTGMAAACAGAGGGCAGATTCAATCCGTTATCTCTAAACAGGTTACAGCAATCCT	241	
Db	304 MetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThrAlaAsnPro	323	
QY	242 AAAGTCACTTTTAAAGATGTCAGAAATAATGATTGATGTCGGAAGGATCATATA-----	295	

Db 684 ValValIleAspAsnArgGlySerCysHisArgGlyLeuIysPheGluGlyAlaPheIys 703
Qy 863 TATAAATGGTCAAATAGAAATTCAGATCAGGTGGAAGGACTCCAATATCTAGCTTCT 922
Db 704 TyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyrLeuAlaSer 723
Qy 923 CGATATGATTCATTGATTCATGATCGTGTGGGCATCCAGGCTGCTCTATGGAGGATAC 982
Db 724 ArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrSerTyrGlyGlyTyr 743
Qy 983 CTCCTCCCTGATGGCATTAAATGCAGAGCTCAGATATCTTCAGGGTTTGCTATTGCTGGGGCC 1042
Db 744 LeuSerIleuMetAlaIleuMetGlnArgSerAspIlePheArgValAlaIleAlaGlyAla 763
Qy 1043 CCAGTCTACTGTGGATCTTCTCTATGATACAGGATACACG 1081
Db 764 ProValThrLeuThrIlePheTyrAspThrGlyTyrThr 776
RESULT 10
AD117085
ID AD117085 standard; protein; 883 AA.
XX
AC AD117085;
XX
DT 15-APR-2004 (first entry)
XX
DE Murine NOVX protein homologue SeqID 621.
XX
KW mouse; murine; NOVX; cardiomyopathy; atherosclerosis; cancer; diabetes;
KW inflammation; autoimmune disorder; allergy; blood disorder;
KW acquired immunodeficiency syndrome; AIDS; obesity; asthma;
KW immunoglobulin (Ig)A nephropathy; cirrhosis; arthritis;
KW Alzheimer's disease; infect.
XX
OS Mus musculus.
XX
FN WO200268649-A2.
XX
PD 06-SEP-2002.
XX
PP 31-JAN-2002; 2002WO-US002785.
XX
PR 31-JAN-2001; 2001US-0265395P.
PR 31-JAN-2001; 2001US-0265412P.
PR 31-JAN-2001; 2001US-0265514P.
PR 31-JAN-2001; 2001US-0265517P.
PR 02-FEB-2001; 2001US-0266406P.
PR 05-FEB-2001; 2001US-0266767P.
PR 07-FEB-2001; 2001US-0266975P.
PR 08-FEB-2001; 2001US-0267057P.
PR 09-FEB-2001; 2001US-0267459P.
PR 15-FEB-2001; 2001US-0267823P.
PR 26-FEB-2001; 2001US-0271664P.
PR 27-FEB-2001; 2001US-0271839P.
PR 27-FEB-2001; 2001US-0271855P.
PR 02-MAR-2001; 2001US-0272788P.
PR 02-MAR-2001; 2001US-0273046P.
PR 14-MAR-2001; 2001US-0275925P.
PR 14-MAR-2001; 2001US-0275947P.
PR 14-MAR-2001; 2001US-0275950P.
PR 14-MAR-2001; 2001US-0275989P.
PR 15-MAR-2001; 2001US-0276448P.
PR 15-MAR-2001; 2001US-0276450P.
PR 16-MAR-2001; 2001US-0276397P.
PR 16-MAR-2001; 2001US-0276768P.
PR 20-MAR-2001; 2001US-0278652P.
PR 26-MAR-2001; 2001US-0278775P.
PR 26-MAR-2001; 2001US-0278778P.
PR 29-MAR-2001; 2001US-0279882P.
PR 29-MAR-2001; 2001US-0279884P.
PR 30-MAR-2001; 2001US-0280147P.
PR 11-APR-2001; 2001US-0282992P.

PR 11-APR-2001; 2001US-0283083P.
PR 20-APR-2001; 2001US-0285133P.
PR 23-APR-2001; 2001US-0285749P.
PR 03-MAY-2001; 2001US-0288327P.
PR 03-MAY-2001; 2001US-0288504P.
PR 29-MAY-2001; 2001US-0294047P.
PR 30-MAY-2001; 2001US-0294473P.
PR 08-JUN-2001; 2001US-0296964P.
PR 18-JUN-2001; 2001US-0298959P.
PR 19-JUN-2001; 2001US-0299324P.
PR 13-AUG-2001; 2001US-0312020P.
PR 16-AUG-2001; 2001US-0312889P.
PR 21-AUG-2001; 2001US-0313390P.
PR 28-AUG-2001; 2001US-0315470P.
PR 31-AUG-2001; 2001US-0316447P.
PR 07-SEP-2001; 2001US-0318115P.
PR 07-SEP-2001; 2001US-0318118P.
PR 12-SEP-2001; 2001US-0318740P.
PR 19-SEP-2001; 2001US-0323379P.
PR 18-OCT-2001; 2001US-0330245P.
PR 18-OCT-2001; 2001US-0330308P.
PR 14-NOV-2001; 2001US-0332701P.
XX
XX (CURA-) CURAGEN CORP.
XX
PI Tchernev VT, Spytek KA, Zerhusen BD, Patturajan M, Shimkets RA;
PI Li L, Gangolli EA, Padigar M, Anderson DW, Rastelli L, Miller CE;
PI Gerlach VL, Taupier RJ, Gusev VY, Colman SD, Wolenc AR, Pena CEA;
PI Furtak K, Grosse WM, Alsobrook JP, Lepley DM, Rieger DK, Burgess CE;
XX
DR WPI; 2002-706998/76.
XX
PT New NOVX polypeptides and nucleic acids, useful for preventing or
PT treating NOVX-associated disorders, e.g. cancer, cardiomyopathy,
PT atherosclerosis, or diabetes, and in chromosome mapping, tissue typing or
PT pharmacogenomics.
XX
PS Disclosure; SEQ ID NO 621; 1498pp; English.
XX
CC This invention relates to a novel nucleic acids, and encoded polypeptides
CC thereof, which have properties related to the stimulation of biochemical
CC or physiological responses in a cell, tissue, organ or organism.
CC Specifically, it refers to the use of biologically active fragments for
CC diagnostic and prognostic assays and furthermore in the treatment of
CC diverse pathological conditions. The present invention describes novel
CC human and murine NOVX proteins, as well as methods to modulate their
CC expression using antisense oligos, ribozymes and peptide nucleic acids.
CC The NOVX polypeptides, polynucleotides and antibodies are useful in
CC treating or preventing NOVX-associated disorders, e.g. cardiomyopathy,
CC atherosclerosis, cancer and diabetes. Furthermore, they may be used in
CC treating or preventing diseases such as inflammation, autoimmune
CC disorders, allergies, blood disorders, acquired immunodeficiency syndrome
CC (AIDS), obesity, asthma, immunoglobulin (Ig)A nephropathy, cirrhosis,
CC arthritis, Alzheimer's disease, infections, stroke, muscular dystrophy
CC and epilepsy. Accordingly, these molecules have many activities including
CC cytotostatic, cardiant, antiinflammatory, immunosuppressive, antiallergic,
CC homeostatic, anti-HIV, antidiabetic, antiarteriosclerotic, anorectic,
CC antiasthmatic, nephrotropic, antiarthritic, hepatotropic,
CC neuroprotective, nootropic, antibacterial, virucide, antiparasitic,
CC relaxant and anticonvulsant. In addition, they are useful in screening
CC assays to identify small molecules that modulate or inhibit, for example,
CC neurogenesis, wound healing and angiogenesis. The nucleic acids are also
CC used as in chromosome mapping, tissue typing, preventive medicine and
CC pharmacogenomics. This polypeptide is a homologue of a human NOVX protein
CC of the invention.
XX
SQ Sequence 883 AA;
Alignment Scores:
Pred. No.: 3.85e-195 Length: 883
Score: 1756.50 Matches: 343
Percent Similarity: 65.7% Conservative: 7

Best Local Similarity: 64.4%		Matches: 10	
Query Match: 90.6%		Indels: 173	
DB:		Gaps: 1	
US-10-825-632-8 (1-1083) x AD117085 (1-883)			
QY	2	GAGAGATGCCAGATCAGCTGAGTCGCTACCTTTGTTCTCCAAAGAAATTTGATAGA	61
DB	245	GlulGluAspProArgSerAlaGlyValAlaThrPheValLeuGlnGluPheAspArg	264
QY	62	TATCTGGCTATTGGTGTCCAAAGCTGAAACAACTCCAGTGGTGGTAAATTCCTT	121
DB	265	TyrSerGlyTyrTrpTrpCysProGlnAlaGluArgThrProSerGlyGlyLysIleLeu	284
QY	122	AGAATTCATATGAAGAAATCATCAATCTGAGGTGGAAATTAATTCATGTTACATCCCT	181
DB	285	ArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisValThrSerPro	304
QY	182	ATGTTGGAACAAAGGAGGCGACATTCATTCCTGCTTATCTTAAACAGGTACAGCAATCCT	241
DB	305	MetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThrAlaAsnPro	324
QY	242	AAAGTCACATTTTAAGATGTCAGAAATAATGATTGATGCTGAAGGAGGATCATATA	295
DB	325	LysValThrPheLysMetSerGluIleValValAspAlaAlaGlyGlyIleAspVal	344
QY	295	-----	295
DB	345	IleAspLysGluLeuValGlnProPheGluIleLeuPheGluGlyValGluTyrIleAla	364
QY	295	-----	295
DB	365	ArgAlaGlyTrpThrProGluGlyLysHisAlaTrpSerIleLeuLeuAspArgSerGln	384
QY	295	-----	295
DB	385	ThrHisLeuGlnIleValLeuIleSerProGluLeuPheIleProValGluAspAla	404
QY	295	-----	295
DB	405	MetAspArgGlnArgLeuIleGluSerValProAspSerValThrProLeuIleIleTyr	424
QY	295	-----	295
DB	425	GluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPheProGlnThr	444
QY	295	-----	295
DB	445	HisGluAspGluIleGluPheIlePheAlaSerGluCysLysThrGlyPheArgHisLeu	464
QY	295	-----	295
DB	465	TyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGlyGlyLeuPro	484
QY	295	-----	295
DB	485	AlaProSerAspPheLysCysProLysGluGluIleThrIleThrSerGlyGluTrp	504
QY	296	-----GTTGATGAAGTCAGAGCTGGTATAT	322
DB	505	GluValLeuGlyArgHisGlySerAsnIleTrpValAspGluAlaArgLysLeuValTyr	524
QY	323	TTTGAGGACCAAGACTCCCTTTAGAGCATCACTGTACGTAGTCAGTTACGTAAAT	382
DB	525	PheGluGlyThrLysAspSerProLeuGluHisLeuTyrValThrSerTyrAlaAsn	544
QY	383	CCTGGAGAGTGACAAAGCTGACTGACCGTGGCTACTCACATTCCTGCTGATCAGTCAG	442
DB	545	ProGlyGluValValArgLeuThrAspArgGlyTyrSerHisSerCysCysLeuSerArg	564
QY	443	CACGTGACTCTTTTATAAGTATAGTAAACAGAGAAGATCCACACTGTGTGCTCCCTT	502
DB	565	HisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeu	584
QY	503	TACAGCTATCAAGTCTCTGAGATGACCCAACTTCTCAAAACAAGAGATTTTGGCCACC	562
DB	585	TyrLysLeuSerSerProGluAspAspProValHisLysThrLysGluPheTrpAlaThr	604
QY	563	ATTTTGATTCACAGAGTCTCTTCTCGACTATATCTCTCCAGAAATTTTCTCTTTGAA	622
DB	605	IleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIlePheSerPheGlu	624
QY	623	AGTACTACTGGATTTTACATTTGATGGGATGCTCTACAAGCCCTCATGATCTACAGCCTGGA	682
DB	625	SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly	644
QY	683	AAGAAATATCTACTGTCTGTTTCATATATGTTGTCCTCAGGTGCAGTTTGGTGAATAAT	742
DB	645	LysLysTyrProThrValLeuPheIleTyrGlyProGlnValGlnLeuValAsnAsn	664
QY	743	CGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTAGCCTCTCTAGGTTATGTGTT	802
DB	665	ArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrValVal	684
QY	803	GTAGTGATACACACAGGGATCCTGTACCCGAGGCTTAAATTTGAAGGCGCTTTAAA	862
DB	685	ValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLys	704
QY	863	TATAAATGCTCAATATGAAATTCAGATCAGGTGGAAGCACTCCATATCTAGCTCT	922
DB	705	TyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyrLeuAlaSer	724
QY	923	CGATATGATTCATTCAGTACTAGATCGTGTGGGCATCCAGGCTGCTCTATGGAGTAC	982
DB	725	GlnTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyrGlyGlyTyr	744
QY	983	CTCTCCCTGATGGCATTAATGCAGAGTCAAGATATCTTCAGGGTTGCTATTGCTGGGGCC	1042
DB	745	LeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAlaGlyAla	764
QY	1043	CCAGTCACTCTGTGGATCTTCTATGATACAGGATACAGC 1081	
DB	765	ProValThrLeuTrpIlePheTyrAspThrGlyTyrThr 777	
RESULT 11			
ABU92026			
ID	ABU92026	standard; protein; 831 AA.	
XX	AC	ABU92026;	
XX	DT	15-JUL-2003 (first entry)	
XX	DE	Human protein modification and maintenance molecule-6 (PMMW-6).	
XX	KW	Human; protein modification and maintenance molecule; PMMW; cancer;	
XX	KW	cell proliferation disorder; atherosclerosis; neurological disorder;	
XX	KW	epilepsy; Huntington's disease; stroke; immune disorder; allergy;	
XX	KW	inflammatory disorder; AIDS; developmental disorder; hypothyroidism;	
XX	KW	Cushing's syndrome; gastrointestinal disorder; epithelial disorder;	
XX	KW	infection; cytostatic; antiarteriosclerotic; anticonvulsant; nootropic;	
XX	KW	neuroprotective; cerebroprotective; anti-HIV; antiallergic; vulnerary;	
XX	OS	antiflammatory; thymomimetic.	
XX	XX	Homo sapiens.	
XX	PN	WO2003031939-A2.	
XX	XX	17-APR-2003.	
XX	PD	11-OCT-2002; 2002WO-US032850.	
XX	XX	12-OCT-2001; 2001US-0329689P.	
XX	XX	25-OCT-2001; 2001US-0335703P.	
PR	PR	09-NOV-2001; 2001US-034887P.	
PR	PR	28-NOV-2001; 2001US-0334145P.	
PR	PR	06-DEC-2001; 2001US-0337451P.	

PR 14-DEC-2001; 2001US-0340584P.
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Rangkumar J, Gorvad AE, Baughn MR, Emerling BM, Yang J, Lee SY;
PI Tran UK, Becha SD, Duggan BM, Lee EA, Griffin JA, Li JX;
PI Sprague WW, Hafalia AJA, Chawla NK, Lehr-Mason PM, Kable AE, Yue H;
PI Marquis JP, Yao MG, Richardson TW, Tang TY, Jin P, Chien D;
PI Bhatia U, Burrill JD, Lee S, Blake JJ, Ho A, Zheng W;
XX
DR WPI; 2003-430274/40.
DR N-PSDB; ACA92421.

XX New human protein modification and maintenance molecules (PMMM), useful
PT for diagnosing, treating and preventing diseases or conditions associated
PT with the aberrant PMMM expression e.g. cancer, atherosclerosis, or
PT infections.

XX Claim 1; Page 242-243; 31pp; English.

XX The present invention relates to the isolation of human protein
CC modification and maintenance molecules (PMMM), and the polynucleotide
CC sequences encoding them. A total of 40 PMMM polypeptides (designated PMMM
CC -1 to PMMM-40) are disclosed. The sequences of the invention are useful
CC for diagnosing a condition or disease associated with the expression of
CC PMMM in a subject, preparing a polyclonal or monoclonal antibody, and
CC generating an expression profile of a sample containing the
CC polynucleotides. The diseases or conditions associated with decreased
CC expression or overexpression of PMMM are cell proliferation disorders
CC (e.g. cancer, atherosclerosis), neurological disorders (e.g. epilepsy,
CC Huntington's disease, stroke), immune/inflammatory disorders, (e.g. AIDS,
CC allergies), developmental disorders (e.g. hypothyroidism, Cushing's
CC syndrome), gastrointestinal or epithelial disorders, and infections. The
CC PMMM polypeptides or their fragments are useful in screening compounds
CC for effectiveness as agonists or antagonists of the polypeptides, or in
CC altering the expression of the target polynucleotide and compounds that
CC specifically bind to, or modulate the activity of the polypeptide.
CC ABU92021-ABU92060 represent the human PMMM polypeptides of the invention

XX Sequence 831 AA;

Alignment Scores:
Pred. No.: 3 616-169 Length: 831
Score: 1534.00 Matches: 309
Percent Similarity: 58.0% Conservative: 0
Best Local Similarity: 58.0% Mismatches: 0
Query Match: 79.2% Indels: 224
DB: 6 Gaps: 2

US-10-825-632-8 (1-1083) x ABU92026 (1-831)

QY 2 GAAGAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAGAAGAATTGTGATAGA 61
DB 244 GUGLUSPALAARGSERALAGLYVALLATHRPHEVALLEUGLUGLUPHEASPARG 263
QY 62 TAATCTGGCTATTGGTGTGTCCAAAAGCTGAAACAACTCCACGTGGTGTAAATTCCT 121
DB 264 Tyr-Ser-Gly-Tyr-Trp-Cys-Pro-Lys-Ala-Glu-Thr-Thr-Pro-Ser-Gly-Gly-Lys-Leu 283
QY 122 AGAATCTATGAAGAAATGATGAATCTGAGGTGGAAATATTATGTTACATCCCT 191
DB 284 Arg-Leu-Tyr-Glu-Glu-Asn-Asp-Glu-Ser-Glu-Val-Glu-Leu-Ile-His-Val-Thr-Ser-Pro 303
QY 182 ATGTTGGAACAGGAGGACAGTTCATTCCTGTTATCTTAAACAGGTACGCAATCCT 241
DB 304 Met-Leu-Glu-Thr-Arg-Ala-Asp-Ser-Phe-Arg-Tyr-Pro-Lys-Thr-Gly-Thr-Ala-Asn-Pro 323
QY 242 AAAGTCACATTTTAAAGATGTCAGAAATAATGATTGATGCTGAAGGAGCATCAT- 295
DB 324 Lys-Val-Thr-Phe-Lys-Met-Ser-Glu-Ile-Met-Ile-Asp-Ala-Glu-Gly-Arg-Ile-Ile-Asp-Val 343
QY 295 ----- 295

DB 344 IleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGluTyrIleAla 363
QY 295 ----- 295
DB 364 ArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAspArgSerGln 383
QY 295 ----- 295
DB 384 ThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGluAspVal 403
QY 295 ----- 295
DB 404 MetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeuIleIleTyr 423
QY 295 ----- 295
DB 424 GluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPheProGlnSer 443
QY 295 ----- 295
DB 444 HisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPheArgHisLeu 463
QY 295 ----- 295
DB 464 TyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGlyGlyLeuPro 483
QY 295 ----- 295
DB 484 AlaProSerAspPheLysCysProIleLysGluIleAlaIleThrSerGlyGluTrp 503
QY 296 -----GTTGTAGAGTCAGAGCTGGTATAT 322
DB 504 GluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgLeuValTyr 523
QY 323 TTTTCAAGGACCAACAGCTCCCTTTAGAGATACCTGTAGTAGTCAGTTACGTAAAT 382
DB 524 PheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSerTyrValAsn 543
QY 383 CCTGGAGAGGTCAAGGCTGACTGACCTGGCTACTCACATTCCTTGTGTCAGTCAG 442
DB 544 ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysIleSerGln 563
QY 443 CACTGTGACTTCTTTATAAGTAACTAGTAAACCAAGAAGATCCACACTGTGTGTCCTT 502
DB 564 HisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeu 583
QY 503 TACAAGCTATCAAGTCTCTGAAAGATGCCCACTTCGCAAAACAAGGAATTTTGGCCACC 562
DB 584 TyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPheTrpAlaThr 603
QY 563 ATTTTGGATTACAGAGCTCTCTTCTGACTATCTCTCCAGAAATTTTCTCTTTGAA 622
DB 604 IleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePheSerPheGlu 623
QY 623 AGTACTACTGGATTACATTGATGGATGCTCAAGCCTCATGATCTACAGCTGGA 682
DB 624 SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly 643
QY 683 AAGAAATATCTACTGTCTGTTTCATATATGTTGGTCTCTCAGGTGCAGTGGTGAATAAT 742
DB 644 LysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeuValAsnAsn 663
QY 743 CGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTAGCCTCTCTAGGTATGTGTT 802
DB 664 ArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrValVal 683
QY 803 GTAGTGATGACCAACAGGGGATCTCTGTCCAGGGGCTTAAATTTGAAGGCGCCTTAAA 862
DB 684 ValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLys 703
QY 863 TATTAATGGTCAATAGAAATTTGACGATCAGGTGGAAGGACTCCATATCTAGCTTCT 922
DB 704 TyrLysMet----- 706


```
QY 289 ----- 289
Db 141 SerGlyLeuProAlaProSerAspPheLysCysProIleLysGluGluLeAlaIle 160
QY 290 -----ATCATAGTTGATGAAGTC 307
Db 161 ThrSerGlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluVal 180
QY 308 AGAAGCTGGTATATTTTGAAGGACCCMAAGACTCCCTTTAGAGCATCACCTGTACGTA 367
Db 181 ArgArgLeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeuTyrVal 200
QY 368 GTCAGTACGTAATCTGGAGAGCTGACAGGCTGACCGTGGCTACTCACATCT 427
Db 201 ValSerTyrValAsnProGlyValThrArgLeuThrAspArgGlyTyrSerHisSer 220
QY 428 TGCTGCATCAGTCAGCACTGTGACTCTTTTATAAGTATAGTAAACCGAAGAATCCA 487
Db 221 CysCysIleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnPro 240
QY 488 CACTGTGTCTCTTTACAGCTATCAAGTCTCAAGATGACCACTTGCACAAACAAG 547
Db 241 HisCysValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLys 260
QY 548 GAATTTTGGGCCACCATTTTGGATTTCAGCAGTCTCTTCTGACTATATCTCTCCAGAA 607
Db 261 GluPheTrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGlu 280
QY 608 ATTTTCTCTTTTGAAGTACTACTGGATTTACATTTGATGGATGCTCTCAAGCTCAT 667
Db 281 IlePheSerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHis 300
QY 668 GAICTACAGCTGGAAAGAAATATCTACTGTCTGTCTTATATATGTGTCTCTCAGGTG 727
Db 301 AspLeuGlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnVal 320
QY 728 CAGTTCGTGAATAATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCTAGCTCT 787
Db 321 GlnLeuValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSer 340
QY 788 CTAGGTATGTGTTGTAGTAGACACACAGGGATCCTGTCTCACCGAGGGCTTAAATTT 847
Db 341 LeuGlyTyrValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPhe 360
QY 848 GAAGGGCCCTTTAAATATATAATGGTCAATAGAAATTGACGATCAGGTGGAGGACTC 907
Db 361 GluGlyAlaPheLysTyrLysMet ----- 368
QY 908 CAATATCTAGCTTCTCGATATGATTTTCATTGACTTTAGATCGTGTGGCATCCACGGCTGG 967
Db 368 ----- 368
QY 968 TCCTATGGAGGATACCTCTCCCTGATGCGATTAAATGCAGAGGTCAAGATATCTTCAGGGTT 1027
Db 369 -----Val 369
QY 1028 GCTATTGCTGGGGCCCGCAGTCACTCTGTGGATCTTCTATGATACAGGATACAG 1081
Db 370 AlaIleAlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThr 387

RESULT 13
AAB93565
ID AAB93565 standard; protein; 632 AA.
XX
AC AAB93565;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human protein sequence SEQ ID NO:12964.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.
XX
```

```
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
XX 28-JUL-2000; 2000EP-00116126.
XX
XX 29-JUL-1999; 99JP-00248036.
XX
PR 27-AUG-1999; 99JP-00300253.
PR
PR 11-JAN-2000; 2000JP-00118776.
PR
PR 02-MAY-2000; 2000JP-00183767.
PR
PR 09-JUN-2000; 2000JP-00241899.
XX
XX (HELI-) HELIX RES INST.
XX
XX Ota T, Isoqai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
XX WPI; 2001-318749/34.
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-
PT length cDNAs defined in the specification, and for the detection and/or
PT diagnosis of the abnormality of the proteins encoded by the full-length
PT cDNAs.
XX
XX Claim 8; SEQ ID NO 12964; 2537pp + Sequence Listing; English.
XX
XX The present invention describes primer sets for synthesizing 5602 full-
XX length cDNAs defined in the specification. Where a primer set comprises:
XX (a) an oligo-dT primer and an oligonucleotide complementary to the
XX complementary strand of a polynucleotide which comprises one of the 5602
XX nucleotide sequences defined in the specification, where the
XX polynucleotide comprises at least 15 nucleotides; or (b) a combination
XX of an oligonucleotide comprising a sequence complementary to the
XX complementary strand of a polynucleotide which comprises a 5'-end
XX sequence and an oligonucleotide comprising a sequence complementary to a
XX polynucleotide which comprises a 3'-end sequence, where the
XX oligonucleotide comprises at least 15 nucleotides and the combination of
XX the 5'-end sequence/3'-end sequence is selected from those defined in the
XX specification. The primer sets can be used in antisense therapy and in
XX gene therapy. The primers are useful for synthesizing polynucleotides,
XX particularly full-length cDNAs. The primers are also useful for the
XX detection and/or diagnosis of the abnormality of the proteins encoded by
XX the full-length cDNAs. The primers allow obtaining of the full-length
XX cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
XX AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
XX represent human amino acid sequences; and AAH13629 to AAH13632 represent
XX oligonucleotides, all of which are used in the exemplification of the
XX present invention
XX
SQ Sequence 632 AA;
XX
Alignment Scores:
Pred. No.: 1,61e-168 Length: 632
Score: 1528.00 Matches: 308
Percent Similarity: 57.8% Conservative: 0
Best Local Similarity: 57.8% Mismatches: 1
Query Match: 78.8% Indels: 224
DB: 4 Gaps: 2
US-10-825-632-8 (1-1083) x AAB93565 (1-632)
QY 2 GAAGAAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAAGAAGAATTGATAGA 61
Db 45 GluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluPheAspArg 64
QY 62 TATTCTGGCTATTGGTGGTGTCTCCAAAGCTGAAACACTCCAGTCGGTGAATAATTTCT 121
Db 65 TyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGlyLysIleLeu 84
QY 122 AGAATTCCTATATGAGAAATGATGATCTGAGGTGGAATATTTCATGTTACATCCCT 181
```

Db 85 ArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisValThrSerPro 104
QY 182 ATGTTGGAACAAGAGGGCAGATTTCATTCCTGTTATCCTTAAACACAGGTACAGCAATCCT 241
Db 105 MetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThrAlaAsnPro 134
QY 242 AAAGTCACCTTTTAAGATGTCAGAAATAAATGATTGCTGTAAGNAGATCATATA 295
Db 125 LysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIleAspVal 144
QY 295 ----- 295
Db 145 IleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGluTyrIleAla 164
QY 295 ----- 295
Db 165 ArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAspArgSerGln 184
QY 295 ----- 295
Db 185 ThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGluAspAspVal 204
QY 295 ----- 295
Db 205 MetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeuIleIleTyr 224
QY 295 ----- 295
Db 225 GluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPheProGlnSer 244
QY 295 ----- 295
Db 245 HisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPheArgHisLeu 264
QY 295 ----- 295
Db 265 TyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGlyGlyLeuPro 284
QY 295 ----- 295
Db 285 AlaProSerAspPheLysCysProIleLysGluGluLeuAlaIleThrSerGlyGluTrp 304
QY 296 -----GTTGATGAAGTCAGAGCGTGATAT 322
Db 305 GluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArgLeuValTyr 324
QY 323 TTGGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTACGTAGTACGTACGTAAAT 382
Db 325 PheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSerTyrValAsn 344
QY 383 CTGGAGAGGTGACAGGCTGACTGACCGTGGCTACTCACATTCCTGCTGCATCAGTCAG 442
Db 345 ProGlyGluValThrArgLeuThrAspArgSerTyrSerHisSerCysLysSerGln 364
QY 443 CACTGTGACTCTTTTATAAGTAAGTATAGTAACAGAGAATCCACACTGTGTGCCCTT 502
Db 365 HisCysAspPheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeu 384
QY 503 TACAAGCTATCAAGTCCTGAAGATGACCCAACTTGCRAACAAAGAAATTTGGGCCACC 562
Db 385 TyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPheIleThrPalatr 404
QY 563 ATTTTGGATTACAGAGGTCCTCTTCCTGACTATACCTCCAGAAATTTCTTTGAA 622
Db 405 IleLeuAspSerAlaGlyProLeuProAspIlyrThrProGluIlePheSerPheGlu 424
QY 623 AGTACTACTGATTTTACATTTGATGGGATGCTCTACAGCTCATGATCTACAGCTCGA 682
Db 425 SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly 444
QY 683 AAGAAATCTCTACTGCTGCTTTCATATATGTTGGTCTCAGGTGAGTTGTTGATAT 742
Db 445 LysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeuValAsnAsn 464

QY 743 CGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCTAGCCTCTCTAGTTTATGTGTT 802
Db 465 ArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrValVal 484
QY 803 GTAGTGATAGACAACAGGGGATCCTGTCCACCGAGGCTTTAAATTTTGAAGGCGCCTTTAAA 862
Db 485 ValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLys 504
QY 863 TATAAAATGGGTCAAAATAGAAATTCAGATCAGGTGGAAGGACTCCAATATCTAGCTTCT 922
Db 505 TyrLysMet ----- 507
QY 923 CGATATGATTTTCATTGACTTTAGATCGTGTGGCACCACGGCTGCTCTATGAGGATAC 982
Db 507 ----- 507
QY 983 CTCTCCCTGATGCATTAAATGCAGAGTTCAGATATCTTCAGGGTTGCTATTGCTGGGGCC 1042
Db 508 -----ValAlaIleAlaGlyAla 513
QY 1043 CCAGTCACTGTGTGGATCTTCTATGATACAGGATACAG 1081
Db 514 ProValThrLeuTrpIlePheTyrAspThrGlyTyrThr 526
RESULT 14
ADR41398
ID ADR41398 standard; protein; 587 AA.
XX ADR41398;
AC ADR41398;
DT 07-OCT-2004 (first entry)
DE Human CD-like molecule HDQES11, SEQ ID NO:197.
XX
KW Human; CD-like molecule; cluster of differentiation; diagnosis;
KW prevention; immune disorder; immunodeficiency; autoimmune disorder;
KW blood-related disorder; haematological disorder; haemostatic disorder;
KW thrombolytic disorder; hyperproliferative disorder; cancer; tumour;
KW apoptotic disorder; cardiovascular disorder; respiratory disorder;
KW angiogenic disorder; neovascularisation; neurological disorder;
KW endocrine disorder; reproductive system disorder; infectious disease;
KW gastrointestinal disorder; drug screening; tissue regeneration;
KW chemotaxis; gene therapy; antibody therapy; drug targeting;
KW chromosome mapping; forensic analysis; immunophenotyping; cytostatic;
KW haemostatic; tranquiliser; vulnery; antiinflammatory; nephrotropic;
KW cardiant; antiallergic; anti-HIV; antirheumatic; antiarthritic;
KW antipsoriatic; immunosuppressive; vasotropic; nootropic;
KW antithyroid; thyromimetic; gynaecological; virucide; hepatotropic;
KW antibacterial; dermatological; chromosome 15q22.
OS Homo sapiens.
XX
PN WO200226930-A2.
XX
PD 04-APR-2002.
XX
PF 25-SEP-2001; 2001WO-US029838.
XX
PR 26-SEP-2000; 2000US-0235484P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Birse CE;
XX
DR WPI; 2002-405050/43.
DR N-PSDB; ADR41222.
XX
PT Novel polynucleotides and polypeptides useful for treating, preventing or
PT ameliorating cardiovascular, renal, neurovascular, and autoimmune
XX disorders.
PS Claim 11; SEQ ID NO 197; 1243pp; English.

XX The invention relates to 167 novel human CD (cluster of differentiation)-
CC like molecules (ADR41388-ADR41563) and to cDNAs encoding them (seqid:1)-
XX
SQ Sequence 587 AA;

Alignment Scores:

Pred. No.: 3,06e-168 Length: 587
Score: 1525.50 Matches: 304
Percent Similarity: 63.5% Conservative: 0
Best Local Similarity: 63.5% Mismatches: 0
Query Match: 78.7% Indels: 175
DB: 5 Gaps: 2

US-10-825-632-8 (1-1083) x ADR41398 (1-587)

QY	164	ATTTCATGTTACATCCCTATGTCGAAACAAGAGGCGCAGATTTCATTCGTTATCCTAAA	223
Db	5	IleHisValThrSerProMetLeuGluThrArgAlaAspSerPheArgTyrProLys	24
QY	224	ACAGGTACAGCAAAATCCCTAAAGTCACCTTTTAAGATGTCAGAAATATGATTCATCCTGAA	283
Db	25	-----ThrAlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGlu	42
QY	284	CGAGGATCAT-----	295
Db	43	GlyArgIleIleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGlu	62
QY	295	-----	295
Db	63	GlyValGluTyrIleAlaArgAlaGlyTyrThrProGluGlyLysTyrAlaTrpSerIle	82
QY	295	-----	295
Db	83	LeuLeuAspArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIle	102
QY	295	-----	295
Db	103	ProValGluAspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerVal	122
QY	295	-----	295
Db	123	ThrProLeuIleIleTyrGluGluThrThrAspIleTyrIleAsnIleHisAspIlePhe	142
QY	295	-----	295
Db	143	HisValPheProGlnSerHisGluGluGluIleGluPheIlePheAlaSerCysLys	162
QY	295	-----	295
Db	163	ThrGlyPheArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArg	182
QY	295	-----	295
Db	183	SerSerGlyGlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAla	202
QY	296	-----GTTGATGAA	304
Db	203	IleThrSerGlyGluTyrGluValLeuGlyArgHisGlySerAsnIleGlnValAspGlu	222
QY	305	GTGAGAGCGTGGTATATTTTGAAGGCACCAAGACTCCCTTTTAGACATCAGCTGTAC	364
Db	223	ValArgArgLeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyr	242
QY	365	GTACTGATACGTAAATCTCGNAGGTCACAGGCTGACTGACCGTGGCTACTACAT	424
Db	243	ValValSerTyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHis	262
QY	425	TCTTGTGTCATCAGTCAGCACTGTGACTCTTCTTATAAGTAAGTATAGTAACACAGAGAAT	484
Db	263	SerCysCysIleSerGlnHisCysAspPheIleSerLysTyrSerAsnGlnLysAsn	282
QY	485	CCACACTGTGTGTCCTTTTACAAGCTATCAAGTCTGTAAGATGACCCCAACTTGCAAAACA	544

Db	283	ProHisCysValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThr	302
QY	545	AAGGAATTTTGGCCACCATTTTGGATTTCAGCAGGTCTCTCTGCTACTATATCTCTCCA	604
Db	303	LysGluPheTyrAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProPro	322
QY	605	GAAATTTCTCTTTTAAAGTACTACTGGATTTTACATTGTATGGATGCTCTCAAGCCT	664
Db	323	GluIlePheSerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysPro	342
QY	665	CATGATCTACAGCCTGGAAAGAAATATCCTACTGCTGCTTCATATATATGCTGCTCAG	724
Db	343	HisAspLeuGlnProGlyLysTyrProThrValLeuPheIleTyrGlyGlyProGln	362
QY	725	GTGCAAGTTGGTCAATAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTAGCC	784
Db	363	ValGlnLeuValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAla	382
QY	785	TCTTAGTATTGTGGTTGTAGTCATAGACAACAGGGGATCCTGTCCAGCGGCTTAAA	844
Db	383	SerLeuGlyTyrValValValIleAspAsnArgGlySerCysHisArgGlyLeuLys	402
QY	845	TTTGAAGCGCCTTTAAATATAATAAATGGTCAATAGAAATTCACGATCAGGTGGAGGA	904
Db	403	PheGluGlyAlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGly	422
QY	905	CTCCAATATCTAGCTTCTCGATATGATTTCATTGACTTAGATCGTGGGCATCCACGCG	964
Db	423	LeuGlnTyrLeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGly	442
QY	965	TGCTCTTAGGAGGATACCTCTCCCTGATGGCATTAATGCAGAGGTGAGATATCTTCAGG	1024
Db	443	TyrSerTyrGlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArg	462
QY	1025	GTTGCTATTGCTGGGGCCCCAGTCACCTCTGTGATCTTCTATGATACAGATACAGC	1081
Db	463	ValAlaIleAlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThr	481
RESULT 15			
AAE14337			
ID	AAE14337	standard; protein; 580 AA.	
AC	AAE14337;		
XX			
DT	07-MAR-2002	(first entry)	
XX			
DE	Human protease PRTS-2	protein.	
KW	Human; protease; PRTS-2; tranquilliser; gene therapy; vaccine; allergy;		
KW	infection; dermatitis; arteriosclerosis; rheumatoid arthritis; hepatitis;		
KW	atherosclerosis; psoriasis; Alzheimer's disease; mental disorder; cancer;		
KW	gastrointestinal disorder; Cushing's syndrome; seizure; glaucoma; stroke;		
KW	epithelial disorder; urticaria; anorexia; trauma; asthma; eczema; nausea;		
KW	hypertension; neurological disorder; Parkinson's disease; drug screening;		
KW	cardiac; cell proliferative disorder; multiple sclerosis; osteoporosis;		
KW	diabetes mellitus; glomerulonephritis; cardiovascular disorder; anaemia;		
KW	autoimmune disorder; inflammatory disorder; myocardial infarction; AIDS;		
KW	developmental disorder; reproductive disorder; infertility; jaundice;		
KW	dementia; acidosis; cataract; gynaecomastia; epilepsy;		
OS	Homo sapiens.		
XX			
PN	WO200183775-A2.		
PD	08-NOV-2001.		
XX			
PF	04-MAY-2001; 2001WO-US014651.		
XX			
PR	04-MAY-2000; 2000US-0202082P.		
PR	11-MAY-2000; 2000US-0203566P.		
PR	17-MAY-2000; 2000US-0205803P.		
PR	25-MAY-2000; 2000US-0207477P.		

```
PR 01-JUN-2000; 2000US-0209402P.
PA (INCY-) INCYTE GENOMICS INC.
XX
XX Deleage AM, Lal P, Hafalia A, Patterson C, Wallia NK, Kearney L;
PI Tribouley CM, Khan FA, Yao MG, Baughn MR, Azimzai Y, Elliott VS;
PI Nguyen DB, Gandhi AR, Yang J, Hernandez R, Policky JL, Lu DAN;
PI Reddy R, Yue H, Tang YT;
XX WPI; 2002-034518/04.
DR N-PSDB; AAD23843.
XX
XX Novel human proteases and polynucleotides encoding the proteases, useful
PT for treating, diagnosing or preventing cell proliferative,
PT cardiovascular, autoimmune/inflammatory, neurological and developmental
PT disorders.
XX
XX Claim 1; Page 120-121; 151pp; English.
XX
XX The invention relates to human proteases (PRTS1-14) and its corresponding
CC cDNA molecules. Human PRTS and its nucleic acid molecule are useful for
CC the diagnosis, treatment and prevention of disorders associated with
CC increased or decreased expression of PRTS. Examples of such disorders
CC include cell proliferative disorders (arteriosclerosis, atherosclerosis,
CC hepatitis, prostatic and cancers); autoimmune/inflammatory disorders
CC (AIDS, Addison's disease, allergy; anaemia, asthma, atopic dermatitis,
CC diabetes mellitus, glomerulonephritis, multiple sclerosis, osteoporosis,
CC trauma, Grave's disease, rheumatoid arthritis, ulcerative colitis, and
CC viral, bacterial, fungal, parasitic, protozoal and helminthic infections)
CC ; cardiovascular disorders (myocardial infarction, ischaemic heart
CC disease and hypertension); neurological disorders (epilepsy, Alzheimer's
CC disease, stroke, mental disorders including mood, anxiety and seasonal
CC affective disorder and prion diseases); gastrointestinal disorders
CC (Crohn's disease, anorexia, nausea, diarrhoea and jaundice); epithelial
CC disorders (contact dermatitis, eczema, acne vulgaris, alopecia, scabies,
CC insect bites and urticaria); reproductive disorder (infertility,
CC disruption of estrous and menstrual cycle and gynaecomastia); and
CC developmental disorders (renal tubular acidosis, Cushing's syndrome,
CC seizure disorders, congenital glaucoma and cataract). PRTS DNA is also in
CC useful in gene therapy. PRTS and its immunogenic fragments are useful for
CC screening libraries of compounds in several drug screening assays. The
CC present sequence is human protease PRTS-2 protein
XX
XX Sequence 580 AA;
SQ
Alignment Scores:
Pred. No.: 1,11e-164 Length: 580
Score: 1495.00 Matches: 306
Percent Similarity: 57.8% Conservative: 0
Best Local Similarity: 57.8% Mismatches: 1
Query Match: 77.1% Indels: 223
DB: Gaps: 2
US-10-825-632-8 (1-1083) x AAE14337 (1-580)
QY 9 ATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAGAGAATTTGATAGATATCTG 68
Db 1 MetProaspGlnLeuGluSerLeuProLeuPheSerLysLysAsnLeuLeuLeuLeu 20
QY 69 GCTATTGGTGTGTCGAAGAGCTGAACAACTCCAGTGGTGGTAAATCTTGAATTC 128
Db 21 AlaileGlyGlyValGlnLysLeuLysGlnLeuProValValLysPheLeuGluPhe 40
QY 129 TATATGAAGAAATGATGAATCTGAGGTGGAATTTATTCATGTTTACATCCCTATGTGG 188
Db 41 TyrMetLysLysMetMetAsnLeuArgTrpLysLeuPheMetLeuHisProLeuCysTrp 60
QY 189 AAACAAGAGG-GCAGATTCAATCCGTTATCCCTAAACAAGGTPACAGCAAAATCCTAAGTC 247
Db 61 LysGlnGlyArgAlaLeuSerPheArgTrpProLysTrpGlyThrAlaAsnProLysVal 80
QY 248 ACTTTTAAGATGTCAGAAATAATGATTGATGCTGAAGAGGATCATATA----- 295
```

```
Db 81 ThrPheLysMetSerGluLeuMetIleAspAlaGluGlyArgIleLeuAspValIleAsp 100
QY 295 -----
Db 101 LysGluLeuLeuGlnProPheGluIleLeuPheGluGlyValGluTrpIleAlaArgAla 120
QY 295 -----
Db 121 GlyTrpThrProGluGlyLysTrpAlaTrpSerIleLeuLeuAspArgSerGlnThrArg 140
QY 295 -----
Db 141 LeuGlnIleValLeuIleSerProGluLeuPheIleProValGluAspValMetGlu 160
QY 295 -----
Db 161 ArgGlnArgLeuIleGluSerValProAspSerValThrProLeuIleIleTrpGluGlu 180
QY 295 -----
Db 181 ThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPheProGlnSerHisGlu 200
QY 295 -----
Db 201 GluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPheArgHisLeuTrpLys 220
QY 295 -----
Db 221 IleThrSerIleLeuLysGluSerLysTrpLysArgSerSerGlyGlyLeuProAlaPro 240
QY 295 -----
Db 241 ThrValThrTrpMetIleThrPheMetArgSerLeuGlyThrProSerCysMetCysVal 260
QY 296 -----GTTGATGAAGTCAGAGCTGGTATATTTTGAAGCACC 334
Db 261 ThrHisIleValGluIleGlnValAspGluValArgLeuValTrpPheGluGlyThr 280
QY 335 AAAGACTCCCTTTAGAGCATCAGCTGTAGTAGTCAGTACGTAAATCTCGAGAGGTG 394
Db 281 LysAspSerProLeuGluHisLeuTrpValValSerValAsnProGlyGluVal 300
QY 395 ACAAGGCTGATGACCGTGGCTACTCAGTTCTGTGTCATCAGTCAGCAGCTGTGCTTC 454
Db 301 ThrArgLeuThrAspArgGlyTrpSerHisSerCysIleSerGlnHisCysAspPhe 320
QY 455 TTTATAAGTATAGTATAGTAAACCAAGAAATCCACACTGTGTGTCCTTTTACAGCTATCA 514
Db 321 PheIleSerLysTrpSerAsnGlnLysAsnProHisCysValSerLeuTrpLysLeuSer 340
QY 515 AGTCTGAAGATGACCCCAACTTCGAAAACAAAGGAATTTTGGGCCACCATTTTGGATTCA 574
Db 341 SerProGluAspAspProThrCysLysThrLysGluPheTrpAlaThrIleLeuAspSer 360
QY 575 GCAGGTCTCTCTCTGACTATCTCTCTCCAGAAATTTCTCTTTTGAAGTACTACTCGA 634
Db 361 AlaGlyProLeuProAspTrpThrProProGluIlePheSerPheGluSerThrThrGly 380
QY 635 TTTACATTGTATGGATGCTCTCAAGCCCTCATGATCTACAGCTCGGAAAGAAATATCTCT 694
Db 381 PheThrLeuTrpGlyMetLeuTrpLysProHisAspLeuGlnProGlyLysLysTrpPro 400
QY 695 ACTGTGCTGTCATATATGCTGCTCTCAGTGCAGTTCGTTGAATTAATCGTTTAAAGA 754
Db 401 ThrValLeuPheIleTrpGlyGlyProGlnValGlnLeuValAsnAsnArgPheLysGly 420
QY 755 GTCAGGTATTTCCGCTTGAATACCTAGCTCTCTAGGTATGCTGTTGCTAGTGTAGTAC 814
Db 421 ValLysTrpPheArgLeuAsnThrLeuAlaSerLeuGlyTrpValValValValIleAsp 440
QY 815 AACAGGGGATCTCTCTACCCAGGCGCTTAAATTTGAAGGCGCTTTTAAATATATAAATGGGT 874
```

```

Db      441 AsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLysTyrllysMet--- 459
QY      875 CAAATAGAAATTGACGATCAGGTGAAGGACTCCAATATCTAGCTTCTCGATATGATTTC 934
Db      459 ----- 459
QY      935 ATTGACTTAGATCGTGTGGGCATCCACGGCTGGTCCCTATGAGGATACCTCTCCCTGATG 994
Db      459 ----- 459
QY      995 GCATTAAATGCAGAGGTGAGATATCTTCAGGGTTGCTATTGCTGGGGCCCCCAGTCACTCTG 1054
Db      460 -----ValAlaIleAlaGlyAlaProValThrLeu 469
QY      1055 TGGATCTTCTATGATACAGGATACAG 1081
Db      470 TrpilePheTyAspThrGlyTyThr 478

```

Search completed: May 2, 2006, 01:03:33
 Job time : 188.411 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: May 2, 2006, 01:03:49 ; Search time 7.34981 Seconds
(without alignments)
2835.522 Million cell updates/sec

Title: US-10-825-632-8
Perfect score: 1938
Sequence: 1 ggaagaagatgccagatcag.....tatgatcacaggatacacgga 1083

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 566832

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ n2p.model -DBV=xlp
-Q=/abs/ABSSWEB.spool/US10825632/runat 01052006 105948 3262/app query.fasta_1
-DB=PIR -QMT=fasta -SUFFIX=n2p.rpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abs02p
-USRR=US10825632.@CGN 1.1.92.@runat 01052006 105948 3262 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DRV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DLEXT=7

Database : PIR 80:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	527.5	27.2	738	2	AB7516 dipeptidyl peptida
2	491.5	25.4	931	2	T32919 hypothetical prote
3	430.5	22.2	741	2	JC5142 x-Pro dipeptidyl-p
4	369	19.0	711	2	S66261 dipeptidyl-peptida
5	315	16.3	766	1	CDHU26 dipeptidyl-aminope
6	312.5	16.1	793	2	T41703 dipeptidyl-peptida
7	309	15.9	760	1	S23752 dipeptidyl-peptida
8	291.5	15.0	792	1	A39914 dipeptidyl-aminope
9	280.5	14.5	818	1	A30107 dipeptidyl-aminope
10	250	12.9	931	2	A49737 dipeptidyl-peptida
11	245	12.6	795	2	F82858 hypothetical prote
12	240	12.4	829	2	T19514 prolyl oligopeptid
13	226.5	11.7	683	2	E87495 dipeptidyl aminope
14	221	11.4	803	2	A41793

15	220	11.4	709	2	B82580 alanyl dipeptidyl
16	219	11.3	779	2	T25173 hypothetical prote
17	219	11.3	799	2	T25174 hypothetical prote
18	205.5	10.6	631	2	H75007 probable acylamino
19	204	10.5	642	2	C71137 hypothetical prote
20	203	10.5	803	2	I68600 dipeptidyl aminope
21	203	10.5	865	2	I54331 dipeptidyl aminope
22	199	10.3	622	2	F71174 hypothetical prote
23	197.5	10.2	743	2	T37700 probable dipeptidyl
24	196	10.1	657	2	E70025 probable acylamino
25	193	10.0	632	2	E75057 peptidase PAB1418
26	189	9.8	536	2	F90299 acylaminoacyl-pept
27	174	9.0	759	2	I38593 fibroblast activat
28	166.5	8.6	674	2	B84381 acylaminoacyl-pept
29	165	8.5	591	2	H72474 probable acylamino
30	162.5	8.4	569	2	S74053 probable acylamino
31	149	7.7	676	2	C97775 acylamino-acid-rel
32	147.5	7.6	667	2	A87711 prolyl oligopeptid
33	147.5	7.6	721	2	T09631 probable acylamino
34	140.5	7.2	591	2	A99444 acylaminoacyl-pept
35	140.5	7.2	606	2	T35378 probable peptide h
36	139.5	7.2	745	2	T33751 hypothetical prote
37	138.5	7.1	598	2	F84199 hypothetical prote
38	137.5	7.1	654	2	AD3183 peptidase [impor
39	134.5	6.9	732	1	JC4655 acylaminoacyl-pept
40	131.5	6.8	732	1	S07624 probable acylamino
41	130.5	6.7	572	2	F72455 acylase and dieste
42	129.5	6.7	555	2	AD1794
43	127.5	6.6	555	2	AB1419
44	126.5	6.5	629	2	T15945 hypothetical prote
45	126.5	6.5	689	1	JN0585 prolyl oligopeptid

ALIGNMENTS

RESULT 1
A87516
dipeptidyl peptidase IV [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C:Accession: A87516
R:NIERMAN, W.C.; FELDBLYUM, T.V.; PAULSEN, I.T.; NELSON, K.E.; EISEN, J.; HEIDELBERG, J.
B.; LAUB, M.T.; DEBOY, R.T.; DODSON, R.J.; DURKIN, A.S.; GWINN, M.L.; HAF, D.H.; KOLON
N, J.; ERMOLAeva, M.; WHITE, O.; SALZBERG, S.L.; SHAPIRO, L.; VENTER, J.C.; FRASER, C.M
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: A87516
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-738 <STO>
A:Cross-references: UNIPROT:Q9A6E0; UNIPARC:UPI000000C7616; GB:AE005673; NID:gl13423647; P:
C:Genetics:
A:Gene: CC2154

Alignment Scores:	1.08e-39	Length:	738
Pred. No.:	527.50	Matches:	137
Score:	41.2%	Conservative:	61
Percent Similarity:	28.5%	Mismatches:	132
Best Local Similarity:	27.2%	Indels:	151
Query Match:	2	Gaps:	13
DB:			
US-10-825-632-8 (1-1083) x A87516 (1-738)			
Qy	5	GAAGATGCCAGTACAGTGGAGTCGCTTCTTCTCAAGAAGAAATTCATAGATAT	64
	:::		
	189	YaspAlaLeuSerPheGlyValAlaGluPheIleValGlnGluGluLeuAspArgPhe	208
Db			
Qy	65	TCTGCTATTGGTGTCTCCAAAAGCTGAACAACATCCCAAGTGGTGGTAAATTTTCAG	124
Db	209	ThrGlyTyrTrpTrpSerProAspGluSer-----Arg	219

125	ATTC	TATATCAAGAAATGATGATCTGAGTGGAAATATTCA	TGTTACATCCCTATG	184
QY				
126	TTGG	AAACAAGGAGCCAGAT		185
QY				
127	TTGG	AAACAAGGAGCCAGAT		186
QY				
128	TTGG	AAACAAGGAGCCAGAT		187
QY				
129	TTGG	AAACAAGGAGCCAGAT		188
QY				
130	TTGG	AAACAAGGAGCCAGAT		189
QY				
131	TTGG	AAACAAGGAGCCAGAT		190
QY				
132	TTGG	AAACAAGGAGCCAGAT		191
QY				
133	TTGG	AAACAAGGAGCCAGAT		192
QY				
134	TTGG	AAACAAGGAGCCAGAT		193
QY				
135	TTGG	AAACAAGGAGCCAGAT		194
QY				
136	TTGG	AAACAAGGAGCCAGAT		195
QY				
137	TTGG	AAACAAGGAGCCAGAT		196
QY				
138	TTGG	AAACAAGGAGCCAGAT		197
QY				
139	TTGG	AAACAAGGAGCCAGAT		198
QY				
140	TTGG	AAACAAGGAGCCAGAT		199
QY				
141	TTGG	AAACAAGGAGCCAGAT		200
QY				
142	TTGG	AAACAAGGAGCCAGAT		201
QY				
143	TTGG	AAACAAGGAGCCAGAT		202
QY				
144	TTGG	AAACAAGGAGCCAGAT		203
QY				
145	TTGG	AAACAAGGAGCCAGAT		204
QY				
146	TTGG	AAACAAGGAGCCAGAT		205
QY				
147	TTGG	AAACAAGGAGCCAGAT		206
QY				
148	TTGG	AAACAAGGAGCCAGAT		207
QY				
149	TTGG	AAACAAGGAGCCAGAT		208
QY				
150	TTGG	AAACAAGGAGCCAGAT		209
QY				
151	TTGG	AAACAAGGAGCCAGAT		210
QY				
152	TTGG	AAACAAGGAGCCAGAT		211
QY				
153	TTGG	AAACAAGGAGCCAGAT		212
QY				
154	TTGG	AAACAAGGAGCCAGAT		213
QY				
155	TTGG	AAACAAGGAGCCAGAT		214
QY				
156	TTGG	AAACAAGGAGCCAGAT		215
QY				
157	TTGG	AAACAAGGAGCCAGAT		216
QY				
158	TTGG	AAACAAGGAGCCAGAT		217
QY				
159	TTGG	AAACAAGGAGCCAGAT		218
QY				
160	TTGG	AAACAAGGAGCCAGAT		219
QY				
161	TTGG	AAACAAGGAGCCAGAT		220
QY				
162	TTGG	AAACAAGGAGCCAGAT		221
QY				
163	TTGG	AAACAAGGAGCCAGAT		222
QY				
164	TTGG	AAACAAGGAGCCAGAT		223
QY				
165	TTGG	AAACAAGGAGCCAGAT		224
QY				
166	TTGG	AAACAAGGAGCCAGAT		225
QY				
167	TTGG	AAACAAGGAGCCAGAT		226
QY				
168	TTGG	AAACAAGGAGCCAGAT		227
QY				
169	TTGG	AAACAAGGAGCCAGAT		228
QY				
170	TTGG	AAACAAGGAGCCAGAT		229
QY				
171	TTGG	AAACAAGGAGCCAGAT		230
QY				
172	TTGG	AAACAAGGAGCCAGAT		231
QY				
173	TTGG	AAACAAGGAGCCAGAT		232
QY				
174	TTGG	AAACAAGGAGCCAGAT		233
QY				
175	TTGG	AAACAAGGAGCCAGAT		234
QY				
176	TTGG	AAACAAGGAGCCAGAT		235
QY				
177				

Db	561	AlaLysPheMetArgAlaLeuAspArgLysLeuGlyThrValGluValGluAspGlnLeu	580
QY	899	GAAGACATCCAATATCTAGCTTCTCGATATGATTTCATTCACCTAGATCTGTGGGCATC	958
Db	581	LeuGlyAlaLysPheLeuAlaSerGln--ProTyrValAspAlaAspLysLeuGlyVal	599
QY	959	CACGGCTGGTCTTATGAGGATACCTCTCCCTGATGGCATTATGCAGAGGTCAGATATC	1018
Db	600	MetGlyTrpSerTyrGlyPheMetAlaLeuMetLeuLeuThrAlaGluAenThrPro	619
QY	1019	TTCAGGTTGCTATTGCTGGGCCCCAGTCACCTCTCTGGATCTTCTATGATACAGATAC	1078
Db	620	PheLysAlaGlyAlaAlaGlyAlaProProThrGluTrpSerLeuTyrAspThrAlaTyr	639
QY	1079	ACG 1081	
Db	640	Thr 640	
RESULT 2			
T32919			
hypoetical protein K02F2.1 - Caenorhabditis elegans			
C:Species: Caenorhabditis elegans			
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004			
R:Accession: T32919			
R:Maggi, L.; Goela, D.			
submitted to the EMBL Data Library, January 1998			
A:Description: The sequence of C. elegans cosmid K02F2.			
A:Reference number: Z21246			
A:Accession: T32919			
A:Status: preliminary; translated from GB/EMBL/DDBJ			
A:Molecule type: DNA			
A:Residues: 1-931 <MAG>			
A:Cross-references: UNIPROT:O44987; UNIPARC:UPI0000076BD8; EMBL:AF043699; PIDN:AAB97564.			
A:Experimental source: strain Bristol N2; clone K02F2			
C:Genetics:			
A:Gene: CESP:K02F2.1			
A:Map position: 1			
A:Introns: 58/3; 82/2; 131/2; 178/2; 275/3; 322/2; 404/2; 441/1; 464/2; 486/3; 528/2; 571/2			
Alignment Scores:			
Pred. No.:	2,08e-36	Length:	931
Score:	491.50	Matches:	145
Percent Similarity:	36.6%	Conservative:	60
Best Local Similarity:	25.9%	Mismatches:	112
Query Match:	25.4%	Indels:	243
DB:	2	Gaps:	19
US-10-825-632-8 (1-1083) x T32919 (1-931)			
QY	23	GGAGTCGCTACCTTCTCTCCAGAGCAATTTGATAGATATTCGGCTATTTGGTGTGT	82
Db	298	GlyValProSerTyrIleValGlnGluLeuGluArgPheGluGlyIleTrpTrp---	316
QY	83	CCAAAGCTGAAACAACACTCCCACTGGTGGTAAATCTTAGAATCTTATGAAGAAAT	142
Db	317	SerGluSerLysThr-----ArgLeuLeuTyrGluHisVal	328
QY	143	GATGAATCTGAGGTGGAATTTATTCATGTT-----ACATCCCT	181
Db	329	AsnGlnGluLysValAlaGluSerGlnPheGlyValAenGlyAspProValAlaPro	348
QY	182	ATGTTGGAACAAGGAGGCAGATTCAATTCGGTTATCTCTAAACAGGTACAGCAATCCT	241
Db	349	Met-----LysTyrProArgAlaGlyThrLysAsnAla	359
QY	242	AAAGTCACCTTTTAAGTG-----	259
Db	360	TyrSerThrLeuArgMetValIleLeuGluAenGlyLysAlaTyrAspValProLeuLys	379
QY	259	-----	259
Db	380	AspGluValIleTyrLysHisCysProPheTyrGluTyrIleThrArgAlaGlyPhePhe	399

```
QY 259 ----- 259
Db 400 SerAspGlyThrThrValTrpValGlnValMetSerArgAspGlnAlaGlnCysSerLeu 419
QY 259 ----- 259
Db 420 LeuLeuIleProTyrThrAspPheLeuLeuProGluGluLeuGlySerIleLeuGlu 439
QY 259 ----- 259
Db 440 AspAsnLeuGlnLeuSerThrAspLeuAsnMetGlyValTrpAspAspLysSerHieGlu 459
QY 259 ----- 259
Db 460 GluThrMetGluLysProProArgGlyLysLeuArgGlyThrValGlnIleHisLysAla 479
QY 259 ----- 259
Db 480 ArgAsnAspTyrTrpIleAsnThrHisAsnAlaIleTyrProLeuLysIleThrAspGlu 499
QY 259 ----- 259
Db 500 GluHisProMetTyrGluPheIleTyrCysLeuGluLysProAsnGlySerCysLeuAla 519
QY 260 ----- TCAGAATAATGATT 274
Db 520 LeuIleSerAlaGluLeuAspGlnAsnGlyTyrCysArgHisThrGluGluLysLeuLeu 539
QY 275 GATGCTGAAGGA ----- AGGATCATAGTTGATGAAGTCAGAGG 313
Db 540 MetAlaGluAsnPheSerIleAsnLysSerMetGlyIleValValAspGluValArgGlu 559
QY 314 CTGGTATATTTGAGGACCAAGACTCCCTTTAGACATCACCCTGTACGTAGTCAGT 373
Db 560 LeuValTyrValAlaAsnGluSerHisProThrGluTrpAsnIle---CysValSer 578
QY 374 TACGTAAATCTCGAGAGGTGACAGGCTGACTGACGCGTGGCTACTCACATTTCTGTGTC 433
Db 579 HisTyrArgThrGlyGlnHisAlaGlnLeuThrGluSerGlyIle----- 593
QY 434 ATCAGTCAGCACGTGCTACTTCTTTATAAGTAAGTATAGTAACCAAGG----- 481
Db 594 -----Cys-----PheLysSerGluArgAlaAsnGlyLysLeuAlaLeuAsp 607
QY 482 AATCCACACATGTGTGTCCTTTTACAAGCTATCA-----AGTCCTGAAGATGACCCAACT 535
Db 608 LeuAspHisGlyPheAlaCysTyrMetThrSerValGlySerProAlaGlu----- 624
QY 536 TGCAAAACAAGGAATTT---TGG----- 556
Db 625 CysArgPheTyrSerPheArgTrpLysGluAsnGluValLeuProSerThrValTyrAla 644
QY 557 GCCACCATTTTGGATTCAGCAGTCTCT---CTTCTGAC-----TATCTCTCCAGAA 607
Db 645 AlaAsnIleThrValSerGlyHisProGlyGlnProAspLeuHisPheAspSerProGlu 664
QY 608 ATTTTCTCTTTGAAAGT---ACTACTGGATTTTACATTGTATGGGATCTCTACAGCCT 664
Db 665 MetIleGluPheGlnSerLysThrGlyLeuMetHisTyrAlaMetIleLeuArgPro 684
QY 665 CATGATCTACAGCTCGAAGAATAATCTCTACTGTGCTTTCATATATGTGGTGGTCTCAG 724
Db 685 SerAsnPheAspProTyrLysLysTyrProValPheHisTyrValTyrGlyProGly 704
QY 725 GTGCAGTTGGTGAATATCGGTTTAAAGGAGTCAAGTATTCGCTTGAATACCTAGCC 784
Db 705 IleGlnIleValHisAsnAspPheSerTrpIleGlnTyrIleArg-----PheCys 721
QY 785 TCTCTAGGTTATGTGGTTGATGATAGACAACAGGGGATCTGTACCCAGGCTTAA 844
Db 722 ArgLeuGlyTyrValValPheIleAspAsnArgGlySerAlaHisArgGlyIleGlu 741
```

```
QY 845 TTTGAAGCGCCTTTAAATATAAATGGGTCAAAATAGAAATGACGATCAGGTGGAAGGA 904
Db 742 PheGluArgHisIleHisLysMetGlyThrValGluValGluAspGlnValGluGly 761
QY 905 CTCGAATATCTAGCTTCTCGATAT---GATTTCAATTGACTTAGATCGTGGGCATCCAC 961
Db 762 LeuGlnMetLeuAlaGluArgThrGlyGlyPheMetAspMetSerArgValValHis 781
QY 962 GCCTGCTCTATGGAGGATACCTCTCCCTGATGGCATTAAATGCAGAGGTGAGATCTTC 1021
Db 782 GlyTrpSerTyrGlyTyrMetAlaLeuGlnMetIleAlaLysHisProAsnIleTyr 801
QY 1022 AGGGTGTGTTATGCTGGGGCCCGCCAGTCACCTCTGTCGATCTCTATCATACAGGATACAG 1081
Db 802 ArgAlaIleAlaGlyAlaValSerAspTrpArgLeuTyrAspThrAlaTyrThr 821

RESULT 3
JC5142
X-Pro dipeptidyl-peptidase (EC 3.4.14.11) - Xanthomonas maltophilia
C:Species: Xanthomonas maltophilia
C>Date: 31-Jan-1997 #sequence_revision 27-Feb-1997 #text_change 09-Jul-2004
C:Accession: JC5142
R:Kobashima, T.; Ito, K.; Yoshimoto, T.
J. Biochem. 120, 1111-1117, 1996
A:Title: Dipeptidyl peptidase IV from Xanthomonas maltophilia: Sequencing and expression
A:Reference number: JC5142; MUID:97164011; PMID:9010758
A:Accession: JC5142
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-741 <KAB>
A:Cross-references: UNIPROT:P95782; UNIPARC:UPI0000085F8A; DDBJ:D83263; NID:g1753196; PI:
C:Comment: This enzyme catalyzes the hydrolytic removal of amino-terminal dipeptidyl res
C:Superfamily: dipeptidyl-peptidase IV
C:Keywords: dipeptidylpeptide hydrolase; membrane bound
F:4-18/Domain: transmembrane #status predicted <TM>
F:610/Active site: Ser #status predicted
F:685/Active site: Asp #status predicted
F:717/Active site: His #status predicted

Alignment Scores:
Pred. No.: 7,09e-31 Length: 741
Score: 430.50 Matches: 125
Percent Similarity: 38.2% Conservative: 57
Best Local Similarity: 26.2% Mismatches: 144
Query Match: 22.2% Indels: 151
DB: 2 Gaps: 13

US-10-825-632-8 (1-1083) x JC5142 (1-741)

QY 23 GGAGTCGCTACTTTGTTCTTCCAGAGAATTTGTAGATATTTCTGGCTATTGCTGGTGT 82
Db 198 GlyIleAlaGluPheValAlaAspGluMetAspArgHisThrGlyTyrTrpAla 217
QY 83 CCAAAAGCTGAACAACTCCCGAGTGGTGTAAATTTCTTAGAATTTCTATATCAAGAAAT 142
Db 218 ProAspSerAla-----IleAlaTyrAlaArgIle 228
QY 143 GATGAATCTGAGGTG-----GAAATATTCACTGTTACATCCCTCATG 184
Db 229 AspGluSerProValProValGlnLysArgTyrGluValTyrAlaAspArgThrAspVal 248
QY 185 TTGAAACAAGAGGCGCAGATTCATTCCTTATCCCTAAACAGGTACAGCAATCCTAAA 244
Db 249 IleGluGln-----ArgTyrProAlaAlaGlyAspAlaAsnValGln 262
QY 245 GTCACCTTTT----- 253
Db 263 ValLysLeuGlyValIleSerProAlaGluGlnAlaGlnThrClnTrpIleAspLeuGly 282
QY 254 AAGATGTCAGAAATA----- 268
Db 283 LysGluGlnAspIleTyrLeuAlaArgValAsnTrpArgAspProGlnHisLeuSerPhe 302
```


Db 381 IleAsnlySValValSerlySLeuAenlleAenThrGlyLyThrGlnLeuSerAsn 400
Qy 410 CGTGGCTACTCAGATCTTCCTGCATCAGTCAGCAGCTGCTTTTATAGTAAGTAT 469
Db 401 AlaGluGlyAsnAsnSerAlaAlaPheSerLyThrPheAsnTyPheIleAsnThrSer 420
Qy 470 AGTAACACAGAAATCCACACTGTGTGCTCCCTTCAAGCTCAAGTCCTGAA----- 523
Db 421 SerThrAlaLySValProThrLySlyTyIleLeuLySAspAlaAenGlyLySAspValLyS 440
Qy 524 -----GATGACCAACTTCACAAACAAAGGAATTTTGGCCACCACTTTTG 568
Db 441 GluLeuGlnAsnAsnAspLeuLeuAenLySLeuLySer----- 454
Qy 569 GATTCAGCAGCTCTCTCTGACTATATCTCTCCAGAAATTTTCTCTTTGAAAGTACT 628
Db 455 -----AspAenPheIleAlaLySgLuPheIleThrIleProAsnAla 468
Qy 629 ACTGGATTTACATTTGATGGATGCTCTACAAGCCTCATGATCTACAGCCTGGAAAGAAA 688
Db 469 AlaGlyAspGlnMetAsnAlaTrpMetIleLySProLySAsnPheAspProAlaLySlyS 488
Qy 689 TATCCTACTGTGCTGTCATATATGTGTGCTCTCAGGTGCAGTTCGTGTAATTCGGTTT 748
Db 489 TyrProValPheMetPheGlnTySerGlyProGlySerGlnGlnValAlaAsnSerTrp 508
Qy 749 AAAGGACTCAAGTATTTCGCTTGATACCTAGCTCTCTAGGTATGTGTTGTAGTG 808
Db 509 AspGlyLyAenGlyIleTrpPheAspMetLeuAlaGlnLySgLyTyLeuValValCys 528
Qy 809 ATAGACAACAGGGGATCTCTGCACCGAGGCTTAAATTTGAAGGCGCTTTAAATATAAAA 868
Db 529 ValAspGlyArgGlyThrGlyPheArgGlyThrLyTyLySlyValThrTyTrpLySAsn 548
Qy 869 ATGGGTCAATAGAAATTCAGATCAGGTGGAAGGACTCCATATCTAGCTTCTCGATAT 928
Db 549 LeuGlyLyTyTrpGluIleGluAspGlnIleThrAlaAlaLySLeuLyGlyAsnGln--- 567
Qy 929 GATTTCACTTACATAGATCTGTGGCATCCAGCTGCTCTCATGGAGGATACCTCTCC 988
Db 568 SerTyValAspLySserArgIleGlyIlePheGlyTrpSerTyGlyGlyTyMetAla 587
Qy 989 CTGATGCATTAAATCAGAGGTCAGATATCTTCAGGGTGTCTATTGTCGGGCCCCAGTC 1048
Db 588 SerLeuAlaMetThrLySgLyAlaAspValPheLySmetGlyIleAlaValAlaProVal 607
Qy 1049 ACTCTGTGGATCTTCATGATACAGGATACAG 1081
Db 608 ThrAsnTrpArgPheTyRAspSerIleTyThr 618

RESULT 5
CDHU26
dipeptidyl-peptidase IV (EC 3.4.14.5) - human
N/Alternate names: cell surface glycoprotein CD26; thymocyte-activating molecule (THAM)
C/Species: Homo sapiens (man)
C/Date: 31-Dec-1993 #sequence revision 23-Aug-1996 #text change 09-Jul-2004
R/Accession: S24313; B61136; S59510; I56154; S59857; S15520
R/Misumi, Y.; Hayashi, Y.; Arakawa, F.; Ikehara, Y.
Biochim. Biophys. Acta 1131, 333-336, 1992
A/Title: Molecular cloning and sequence analysis of human dipeptidyl peptidase IV, a ser
A/Reference number: S24313; MUID:92329551; PMID:1352704
A/Accession: S24313
A/Molecule type: mRNA
A/Residues: 1-6,'I', 8-766 <MIS>
A/Cross-references: UNIPARC:UPI000016A80; EMBL:X60708; NID:g35335; PIDN
R/Darmoul, D.; Lacasa, M.; Baricault, L.; Marguet, C.; Trotot, P.; Barbat, A.
J. Biol. Chem. 267, 4824-4833, 1992
A/Title: Dipeptidyl peptidase IV (CD 26) gene expression in enterocyte-like colon cancer
IV mRNA levels during cell differentiation.
A/Reference number: A42408; MUID:92165847; PMID:1347043
A/Accession: B42408
A/Molecule type: mRNA
A/Residues: 1-5,'R',7-436,'S',438-556,'I',558-662,'E',664-766 <DAR1>

A/Cross-references: UNIPARC:UPI0000052ACB; GB:M80536; NID:g181569; PIDN:AAAS2308.1; PID:
A/Experimental source: intestine
A/Note: this sequence corresponds with the author's translation
A/Accession: A42408
A/Molecule type: mRNA
A/Residues: 1-5,'R',7-436,'S',438-556,'I',558-662,'E',664-711,'G',713-766 <DAR2>
A/Cross-references: UNIPARC:UPI0000172A2B; GB:M80536; NID:g181569
A/Note: sequence extracted from NCBI backbone (NCBIN:83986, NCBIP:83988); this sequence
R/Gorvel, J.P.; Ferrero, A.; Chambraud, L.; Rigal, A.; Bonicel, J.; Maroux, S.
Gastroenterology 101, 618-625, 1991
A/Title: Expression of sucrose-isomaltase and dipeptidylpeptidase IV in human small inte
A/Reference number: A61136; MUID:91317403; PMID:1677636
A/Accession: B61136
A/Molecule type: protein
A/Residues: 1-15,'X',17-22 <GOR>
A/Cross-references: UNIPARC:UPI0000172A2C
R/Boehm, S.K.; Gum Jr., J.R.; Erickson, R.H.; Hicks, J.W.; Kim, Y.S.
Biochem. J. 311, 835-843, 1995
A/Title: Human dipeptidyl peptidase IV gene promoter: tissue-specific regulation from a
A/Reference number: S59510; MUID:96067599; PMID:7487939
A/Accession: S59510
A/Molecule type: DNA
A/Residues: 1-31 <BOE>
A/Cross-references: UNIPARC:UPI000016B4A6; GB:S79876; NID:g1195574; PIDN:AAB35614.1; PID
R/Tanaka, T.; Camerini, D.; Seed, B.; Torimoto, Y.; Dang, N.H.; Kameoka, J.; Dahlberg, H.
J. Immunol. 149, 481-486, 1992
A/Title: Cloning and functional expression of the T cell activation antigen CD26.
A/Reference number: I56154; MUID:92325476; PMID:1352530
A/Accession: I56154
A/Status: translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-436,'S',438-766 <TAN>
A/Cross-references: UNIPARC:UPI000004F7BF; GB:M74777; NID:g180082; PIDN:AAAS1943.1; PID:
R/Abbott, C.A.; Baker, E.; Sutherland, G.R.; McCaughan, G.W.
Immunogenetics 40, 331-338, 1994
A/Title: Genomic organization, exact localization, and tissue expression of the human CD
A/Reference number: S59857; MUID:95012454; PMID:7927537
A/Accession: S59857
A/Molecule type: DNA
A/Residues: 1-436,'S',438-766 <ABB>
A/Cross-references: UNIPARC:UPI000004F7BF; EMBL:UI3734
C/Genetics:
A/Gene: GDB:DPP4
A/Cross-references: GDB:125239; OMIM:102720
A/Map position: 2q24.3-2q24.3
A/Introns: 2/3; 32/1; 95/3; 122/3; 140/2; 164/3; 205/1; 258/3; 296/2; 341/3; 356/3
C/Supfamily: dipeptidyl-peptidase IV
C/Keywords: dipeptidylpeptide hydrolase; glycoprotein; homodimer; proteinase; transmembr
F/1-6/Domain: intracellular #status predicted <INT>
F/7-28/Domain: transmembrane #status predicted <TN>
F/29-766/Domain: extracellular #status predicted <EXT>
F/85.92,150,219,229,281,321,520,685/Binding site: carbohydrate (Asn) (covalent) #status
F/630,708,740/Active site: Ser, Asp, His #status predicted
Alignment Scores:
Pred. No.: 2,25e-20 Length: 766
Score: 315.00 Matches: 120
Percent Similarity: 34.9% Conservative: 59
Best Local Similarity: 23.4% Mismatches: 146
Query Match: 16.3% Indels: 188
DB: 1 Gaps: 22
US-10-825-632-8 (1-1083) x CDHU26 (1-766)
Qy 2 GAAGAAGATGCCAGATCAGTCAGTCGCTACCTGTTTCTCCAAAGAA---TTTGAT 58
Db 190 LysGluAspIleIleTyRAsnGlyIleThrAspTrpValTyRGlulGluValPheSer 209
Qy 59 AGATATTCCTGCTATTGGTGGTGTCTCCAAAGCTGAAACCACTCCAGTCGTGTGTAATAATT 118
Db 210 AlaTySerAlaLeuTrpTrpSerProAsnGlyThrPhe----- 222
Qy 119 CTTAGAAATCTATATGAAAGAAATGATGAATCTGAGTGGAAATATTTCAT----- 169

Db 240 -----LeuGlnTyrProLysThrValTrpIleProTyrPro 251
QY 221 AAAACAGGTACAGCAATCTCTAAAGTCACCTTT----- 253
Db 252 LysAlaGlyAlaValAsnProThrValLysPhePheIleValAsnIleAspSerLeuSer 271
QY 253 ----- 253
Db 272 SerSerSerSerAlaIleProIleGlnIleProAlaProAlaSerValAlaAraGlyAsp 291
QY 254 ---AAGATGTCAGAAATAATGATGTGCTGAAGGAAGGATCATAGTTGATGAAGTCAGA 310
Db 292 HisTyrLeuCysAspValValTrpAlaThrGluGluAraGlySerLeuGlnTrpLeuAra 311
QY 311 AGCTGTGTATTTTGAAGC-----ACCAAA 337
Db 312 ArgIleGlnAsnTyrSerValMetAlaIleCysAspTyrAspLysIleAsnLeuThrTrp 331
QY 338 GACTCCCTTTAGAG---CATCACCTGTACGTA-----GTCAAGTTACGTA 379
Db 332 AsnCysProSerGluGlnGlnHisValGluMetSerThrGlyTrpValGlyArgPhe 351
QY 380 AATCCTGGAGGTGACAGGCTGACTGAC----- 409
Db 352 ArgProAlaGluProHisPheThrSerAspGlySerSerPheTyrLysIleIleSerAsp 371
QY 410 ---CGTGGCTACTCACATCTCTGCTGCATC-----AGTCAGCACTGTGACTTC--- 454
Db 372 LysAspGlyTyrLysHisIleCysHisPheProLysAspLysLysAspCysThrPheIle 391
QY 455 -----TTTATAGTAAG 466
Db 392 ThrLysGlyAlaTrpGluValIleSerIleGluAlaLeuThrSerAspTyrLeuTyr 411
QY 467 TATAGTAACAG-----RAGATCCACACTGTGTCTCCCTTACAG----- 508
Db 412 IleSerAsnGlnTyrLysGluMetProGlyGlyArgAsnLeuTyrLysIleGlnLeuThr 431
QY 509 ---CTATCAAGTCTGAAGATGCCCACTTGCAAA----- 541
Db 432 AspHisThrAsnValLysCysLeuSerCysAspLeuAsnProGluAraCysGlnTyr 451
QY 542 -----ACAAGGAAATTTGGCCACCATTTTGGATTCAGCAGGTCT---CTT 586
Db 452 AlaValSerPheSerLysGluAlaLysTyrTyrGlnLeuGlyCysTrpGlyProGlyLeu 471
QY 587 CQTGACTATACT----- 598
Db 472 ProLeuTyrThrLeuHisArgSerThrAspHisLysGluLeuAraGValLeuGluAspAsn 491
QY 599 -----CTCCAGAAATTTTCTCTTTT--- 619
Db 492 SerAlaLeuAspArgMetLeuGlnAspValGlnMetProSerLysLysLeuAspPheIle 511
QY 620 ---GAAAGTACTACTGATTTACATTTGATGGGATGCTCTCAAGCCCTCATGATCTACAG 676
Db 512 ValLeuAsnGluThrArgPhe---trpTyrGlnMetIleLeuProProHis---PheAsp 529
QY 677 CCTGGAAGAAATATCCTACTGTGCTGTTCATATATGTTGGTGTCTCAGGTGCAGTTGGTG 736
Db 530 LysSerLysLysTyrProLeuLeuLeuAspValTyrAlaGlyProCysSerGln----- 547
QY 737 AATAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAAT-----ACCTAGCCTCT 787
Db 548 -----LysAlaAspAlaSerPheAraGLeuAsnTrpAlaThrTyrLeuAlaSer 563
QY 788 CTAGGTTATGTGTTGTAGTG---ATAGACACAGGGATCCTGTCCAGGAGGCTTAAA 844
Db 564 ThrGluAsnIleIleValAlaSerPheAspGlyArgGlySerGlyTyrGlnGlyAspLys 583
QY 845 TTTGMAAGGCCCTTTAAATATAAATGGGTCAAAATAGAAATTTGACGATCAGGTGGAGGA 904
Db 584 IleMetHisAlaIleAsnArgLeuGlyThrLeuGluValGluAraGlnIleGluAla 603

QY 905 CTCCAATATCTAGCTTCTCGATATGATTTTCATTGACTTACATGCTGTGGCCATCCACGCC 964
Db 604 AlaArgGlnPheVal---LysMetGlyPheValAspSerLysArgValAlaIleTrpGly 622
QY 965 TGGTCTCTATGAGGATACCTCTCCCTGATGGCATTAAATGCAGAGGTTCAGATATCTTCAGG 1024
Db 623 TrpSerTyrGlyTyrValThrSerMetValLeuGlySerGlySerGlyValPheLys 642
QY 1025 GTTCTATTTGCTGGGGCCCGAGTCACCTCTGTGGATCTTCTATGATACATACACG 1081
Db 643 CysGlyIleAlaValAlaProValSerArgTrpGluTyrTyrAspSerValTyrThr 661
RESULT 8
A39914
dipeptidyl-peptidase IV (EC 3.4.14.5), membrane-bound form precursor - rat
N/Alternate names: Gp110; membrane glycoprotein 110K; OX-61
N/Contains: dipeptidyl-peptidase IV, soluble form
C/Species: Rattus norvegicus (Norway rat)
C/Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text change 09-Jul-2004
C/Accession: A39914; A33315; B33315; A60730; A42203; S38949; A31781
R/Hong, W.; Doyle, D.
Proc. Natl. Acad. Sci. U.S.A. 84, 7962-7966, 1987
A/Title: cDNA cloning for a bile canalicular domain-specific membrane glycoprotein of ra
A/Reference number: A39914; MUID:88068516; PMID:3479775
A/Accession: A39914
A/Molecule type: mRNA
A/Residues: 1-792 <HON>
A/Cross-references: UNIPROT:P14740; UNIPARC:UPI000017098A; GB:J02997; NID:G204463; PIDN:
R/Ogata, S.; Misumi, Y.; Ikehara, Y.
J. Biol. Chem. 264, 3596-3601, 1989
A/Title: Primary structure of rat liver dipeptidyl peptidase IV deduced from its cDNA an
A/Reference number: A33315; MUID:89123496; PMID:2563382
A/Accession: A33315
A/Molecule type: mRNA
A/Residues: 1-37; A'39-182; I', 184-331; T', 333-351; C', 353-393; V', 395-561; L', 563-623;
A/Cross-references: UNIPARC:UPI0000129841; GB:J04591; NID:G203973; PIDN:AAA41036.1; FID:
A/Note: the authors translated the codon GCG for residue 38 as Arg, ACC for residue 332
A/Accession: B33315
A/Molecule type: protein
A/Residues: 1-20; 35-54; 427-443; 505-509; 511-520; 530-538; 593-600; 602-608; 618-627 <HO2>
A/Cross-references: UNIPARC:UPI0000172A37
A36; UNIPARC:UPI0000172A37
R/McCaughan, G.W.; Wickson, J.E.; Creswick, P.F.; Gorrell, M.D.
Hepatology 11, 534-544, 1990
A/Title: Identification of the bile canalicular cell surface molecule GP110 as the ectop
quence.
A/Reference number: A60730; MUID:90228896; PMID:1970322
A/Accession: A60730
A/Molecule type: protein
A/Residues: 28-47; XX', 50-53; 55-58 <MCC>
A/Cross-references: UNIPARC:UPI0000172A38
R/Ogata, S.; Misumi, Y.; Teuji, E.; Takami, N.; Oda, K.; Ikehara, Y.
Biochemistry 31, 2582-2587, 1992
A/Title: Identification of the active site residues in dipeptidyl peptidase IV by affini
A/Reference number: A42203; MUID:92190188; PMID:1347701
A/Accession: A42203
A/Molecule type: protein
A/Residues: 'R', 625-630; 'X', 632-648 <OG2>
A/Cross-references: UNIPARC:UPI0000172A39
R/Iwaki-Egawa, S.; Watanabe, Y.; Fujimoto, Y.
Biol. Chem. Hoppe-Seyler 374, 973-975, 1993
A/Title: N-terminal amino acid sequence of the 60-kDa protein of rat kidney dipeptidyl p
A/Reference number: S38949; MUID:94128239; PMID:7905271
A/Accession: S38949
A/Status: preliminary
A/Molecule type: protein
A/Residues: 281-302 <IWA>
A/Cross-references: UNIPARC:UPI0000172A3A
R/Hong, W.; Doyle, D.
J. Biol. Chem. 263, 16892-16898, 1988
A/Title: Membrane orientation of rat gp110 as studied by in vitro translation.
A/Reference number: A31781; MUID:89034185; PMID:3182821

Db	404	GluValIleSerIleGluAlaLeuThrSøAspTyrLeuTyrTyrIleSerAsnGlnTyr	423
QY	479	--AAGAATCCACACTGTGTGTCCTCTTTACAAG-	508
Db	424	LysGluMetProGlyGlyArgAsnLeuTyrIleGlnLeuThrAspHisThrAsnLys	443
QY	509	-----CTATCAAGTCCTGAGATGACCCCACTTGCAAA-ACA	544
Db	444	LysCysLeuSerCysAspLeuAsnProGluArgCysGlnTyrTyrSerValSerLeuSer	463
QY	545	AAGAAATTTGGGCCACCATTTTGGATTGACAGAGTCTCT--CTTCCTGACTTACT---	598
Db	464	LysGluAlaIyStyTyrGlnLeuGlyCysArgGlyProGlyLeuProLeuTyrThrLeu	483
QY	599	-----CCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGATTACA---	640

[illegible]

Db QY Db QY Db QY Db QY Db QY Db QY Db QY Db QY Db QY

543 LeuLeulleAspValTyAlaGlyProCysSerGln-----Lysala 556
755 GTCAAGATATTCCGGTTGAAT-----ACCTAGCCTCTAGGTATTATCGTGGTTGA 805
557 AspAlaAlaPheA-gPheAsnTrpAlaThrTyLeuAlaSerThrGluAsnIleIleVal 576
806 GNG---ATAGACAACAGGGGATCTGTGCACCAGGGGCTTAATAATTGAAGGCCGCTTTAA 862
577 AlaSerPheAspGlyArgGlySerGlyTyTGlnGlyAspLysIleMethHisAlaIleAsn 596
863 TATAAAATGGGTCAAATAGAAAATTGACGATCAGGTGGAGGACTC---CAATATCTTAGCT 919
597 LysArgLeuGlyThrLeuGluValGluAspGlnIleGluAlaalaargGlnPheLeu---- 615
920 TCTCGATATGATTCATTGACTTAGATCGTGTGGGATCCACGGCTGGCTCATATGGAGA 979
616 ---LysMetGlyPheValAspSerLysGlnValAlaIleTrpGlyTrpSerTyrglyGly 634
980 TACCTCTCCCTGATGGCATTAATGACAGAGGTGCAGATATCTTCAGGTTTGCATTGCTGGG 103
635 TyrValThrSerMetValLeuGlySerGlyValPheLysCysGlyIleAlaVal 654
1040 GCCCAGTCACCTCTGTGATCTTCATGATACAGGATACACG 1081
655 AlaProValSerArgTrpGluTyTyraSpSerValTyThr 668

RESULT 9
A30107 dipeptidyl aminopeptidase B (EC 3.4.14.-) - yeast (*Saccharomyces cerevisiae*)
N:Alternate names: protein YHR028c
C:Species: *Saccharomyces cerevisiae*
C>Date: 07-Jun-1990 #sequence_revision 30-May-1997 #text_change 09-Jul-2004
C:Accession: S46780; A30107
R;Du, Z.
submitted to the EMBL Data Library, June 1994
A:Description: The sequence of *S. cerevisiae* cosmid 8082.
A:Reference number: S46773
A:Accession: S46780.
A:Molecule type: DNA
A:Residues: 1-818 <DUZ>
A:Cross-references: UNIPROT:P18962; UNIPARC:UPI00000031A5F; EMBL:U10399; NID:95
R:Roberts, C.J.; Pohlig, G.; Rothman, J.H.; Stevens, T.H.
J. Cell Biol. 108, 1363-1373, 1989
A>Title: Structure, biosynthesis, and localization of dipeptidyl aminopeptidase

THIS PAGE BLANK (USPTO)

A;Reference number: A30107; MUID:89174971; PMID:2647766
A;Accession: A30107

A;Accession: A30107
A;Molecule type: DN

A; Residues: 1-82, 'H'

A; Cross-references: 1
A; Note: the authors

A/NOTE: the authors translated the codon ACC for residue C/Genetics:

A;Gene: SGD:DAP2, STE13; MIPS:YHR028C

A;Cross-references: SGD:S0001070; MIPS:YHR028C
A;Map position: 8R

C; Superfamily: dipeptidyl-peptidase I

C; Keywords: dipeptidylpeptide hydrolase; glycoprotein; peptidase; peptidase IV
E: 3.0-4.5/None

F;30-45/Domain: transmembrane #status predicted
F;63,79,110,139,392,421/Binding site: carbohydrate

1,33,73,110,139,392,421/Binding site: C

Alignment Scores:

Pred. No.: 3.11e-17
Score: 280.50

Percent Similarity: 31.7%

```
Best Local Similarity: 22.0%
Query Match: 14.5%
```

```

Query Matchn:
DB:
14.58
1

```

4

US-10-825-632-8 (1-1083) x A30107 (1-81

QY 23 GGAGTCGCTACCTTTGTTCTCTCCAAG

2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525
 526

234 GlyLysProAspTrpValTyrGluG

QY 80 TGTCCAAAGCTGAAACAACTCCAGT

- - - - -
 |||||
 - - - - -
 : : : : :
 254 GCGGCTGGTGGAGAAATCTAGAAATCTATATGAAGAA 135
 b

Ub 254 SerProThrGlyAsp-----TyrLeuAlaPheLeuLys 264

140 AATGATGAATCTGAGGTG--GAAATTATTCATGTTACATCCCTATGTTGGAAACACACG 186

265 T T A C C G T T G C A T C C C C C T A T G T T G G A A C A A A G G 198

265 IleAspGluSerGluValGlyGluPheIle-----IleProTyrTyrValGlnAsp 281

197 AGGCAGAT-----TCATTCCGTTATCCTAAACAGGTACAGCAAAAT 238

282 GCTTGTGTTTGTTGGTCTTCAT
|||
CGTTCCTGTTATCCCTAAACAGGTACAGCAAT 238

282 GluLysAspIleTyrProGluMetArgSerIleLysTyrProLysSerGlyThrProAsn 301

239 CCT-----241

302 PROHIBITION AGAINST EMPLOYMENT OF
 --- 241

302 ProHisAlaGluLeuTrpValTyrSerMetLysAspGlyThrSerPheHisProArgIle 321

241 ----- 241

b
322 SERGI VASILVICHENCO
----- 241

322 SerGlyAsnLysLysAspGlySerLeuLeuIleThrGluValThrTrpValGlyAsnGly 341

242 -----AAAGTCACTTTTAAGATGTCAGAAATA-----ATGATTGAT 277

b
342 AspValLeuValIvstThrThrAsnArgSerGluValLys
:::|:::|
||| |||

342 ASNVALDEUVALYLSTHRTHRASPARGSSERASPILLELEUTHRVLPHELEUILEASP 361

277 ----- 277

b
362 ThrIleAlaLysThrSerAsnValValArgAsnGluSerSerValGluCyl-MetMetCyl-

ValArgAsnGluSerAsnGlyGlyTrpTrpGlu 381

y ----- 277

b

382 IleThrHisAsnThrLeupheIleProAlaAsnGluThrpheAsnArgpro;CAGACG... 403

100SASG.....LLEDFLELLEFFOALDASNGIUTHPheAspArgProHisAsnGly 401

277 ----- 277

A;Reference number: S45451; MUID:95066382; PMID:7975897		Db		533 LysGluProIleGlnLeuThrLysGlyAsnTrpGluValThrGlyAsnGlyValGly 552	
A;Accession: S45451		Qy		296 GTTGATGAAGTCAGAGGCTGGTATATTTTGAAGGACCAAGACTCCCTTTAGAGCAT 355	
A;Molecule type: DNA		Db		553 TyrGluTyrGluThrAspThrIlePhePheThrAlaAsnGluIleGlyValMetSerGln 572	
A;Residues: 1-931 <ANN>		Qy		356 CACCTGTACCTAGTCAGTTAGTA----- 379	
A;Cross-references: UNIPARC:UPI0000136060; EMBL:L21944; NID:g347196; PIDN:AAA35119.1; R:Gallison, F.; Dujon, B.		Db		573 HisLeuTyrSerIleSerLeuThrAspSerThrThrGlnAsnThrPheGlnSerLeuGln 592	
submitted to the EMBL Data Library, October 1995		Qy		380 AATCTCGGAGGAGTCAGACGCTCAGCCGCTGCTACTCACAATTCCTTGTGCTGCTACT 439	
A;Description: Sequence and analysis of a 33 kb fragment from the right arm of chromosome		Db		593 AsnPro-----SerAspLysTyrAspPheTyrAspPheGluLeuSer 606	
A;Reference number: S60938		Qy		440 CAGCACTGTGACTCTCTTTTATAAGTATAGTAACACGAGGAATCCA----- 487	
A;Accession: S60946		Db		607 SerSerAlaArgTyrAlaIleSerLysLysLeuGlyProAspThrProIleLysValAla 626	
A;Molecule type: DNA		Qy		488 -----CACTGTGTGCTCCCTTTTCAAGCTA 511	
A;Residues: 1-931 <GAL>		Db		627 GlyProLeuThrArgValLeuAsnValAlaGluIleHisAspSerSerIleLeuGlnLeu 646	
A;Cross-references: UNIPARC:UPI0000136060; EMBL:X92441; NID:g1050762; PIDN:CAA63182.1; R:Boyer, J.; Fairhead, C.; Gaillon, L.; Gallison, F.; Michaux, G.; Thierry, A.; Dujon, B.		Qy		512 TCAAGTCTCTGAGATGACCCCAACTTGCAAAACAAAGGAATTTTGGGCCACCACTTTTGGAT 571	
submitted to the Protein Sequence Database, July 1996		Db		647 ThrLysAspGlu-----LysPheLysGlu----- 654	
A;Reference number: S71713; MUID:96437977; PMID:8840505		Qy		572 TCAGCAGGTCCTCTCTCTGACTATATCTCCCAAGAAATTTTCTCTTTTGAAGTACTACT 631	
A;Accession: S71721		Db		655 -----LysIleLysAsnTyrAspLeuPro---IleThrSerTyrLysThrMetVal 670	
A;Status: nucleic acid sequence not shown; translation not shown		Qy		632 -----GGATTTTACATTGTATGGGATGCTCTACAGGCTCATGATCTACAGCCTGGA 682	
A;Molecule type: DNA		Db		671 LeuAspAspGlyValGluIleAsnTyrIleGluIleLysProAlaAsnLeuAsnProLys 690	
A;Residues: 1-931 <GAW>		Qy		683 AGAAATATCTCTACTGCTGCTTCTATATATGCTGCTCTCAGGTGCGATTTGGTAATAAT 742	
A;Cross-references: UNIPARC:UPI0000136060; EMBL:X92441; NID:g1050762; PIDN:CAA63182.1; R:Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1995		Db		691 LysLysTyrProIleLeuValAsnIleTyrGlyProGlySerGlnThrPheThrThr 710	
C;Genetics:		Qy		743 CGGTTTAAAGGAGTCAAGTATTTCGCTTGAATACCTAGCTCTCTAGGTTTGTGGTT 802	
A;Gene: SGD:STE13; YC11		Db		711 -----LysSerSerLeuAlaPheGluGlnAlaValSerGlyLeuAspValIleVal 728	
A;Cross-references: SGD:S0005745; MIPS:YOR219c		Qy		803 GTAGTCATAGACACACAGGGATCTCTGCACCGAGGCTTAAATTTGAAGGCGCTTTAAA 862	
A;Map position: 15R		Db		729 LeuGlnIleGluProArgGlyThrGlyGlyLysGlyTyrSerPheArgSerTrpAlaArg 748	
C;Function:		Qy		863 TATAAATGGTCAATAGAAATTCAGCATCAGGTGGAA---GGACTCCAATATCTAGCT 919	
A;Description: involved in processing of alpha-factor prepropheromone		Db		749 GluLysLeuGlyTyrTrpGluProArgAspIleThrGluValThrLysLysPheIleGln 768	
F;Keywords: dipeptidylpeptide hydrolase; glycoprotein; transmembrane protein		Qy		920 TCTCGATATGATTTCAATTCAGCTTAGATCGTGTGGGCATCCACGCTGCTCTCATGGAGGA 979	
F;113-150/Domain: transmembrane #status predicted <TM>		Db		769 ArgAsnSerGlnHisIleAspGluSerLysIleAlaIleTrpGlyTyrSerTyrGlyGly 788	
F;377/Binding site: carbohydrate (Asn) (covalent) #status predicted		Qy		980 TACCTCTCCCTGATGCGATTA---ATCGCAGAGTCAGATATCTTCAGGGTTGCTATTGCT 1036	
F;785,863,896/Active site: Ser, Asp, His #status predicted		Db		789 PheThrSerLeuLysThrValGluLeuAspAsnGlyAspThrPheLysTyrAlaMetAla 808	
Alignment Scores:		Qy		1037 GGGGCCCCAGCTCACTCTGTGGATCTTCTTATGATACAGGATACAG 1081	
Pred. No.:	1.87e-14	Db		809 ValAlaProValThrAsnTrpThrLeuTyrAspSerValTyrThr 823	
Score:	250.00	Qy		RESULT 11	
Best Local Similarity:	24.3%	Db		F82858	
Query Match:	12.9%	Qy		dipeptidyl-peptidase XF0015 [imported] - Xylella fastidiosa (strain 9a5c)	
DB:	2	Db		C;Species: Xylella fastidiosa	
US-10-825-632-8 (1-1083) x A49737 (1-931)		Qy		C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004	
		Db		C;Accession: F82858	
		Qy		R;Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen	
		Db		Nature 406, 151-157, 2000	
		Qy		A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.	
		Db		A;Reference number: A82515; MUID:20365717; PMID:10910347	

Db	283	ArgGluValAsnArgIlePheLeuProIlytyrThrAspAsp	296
Qy	200	GCAGATTCAATC	241
Db	297	---AspSerTyrValGluTyrPheGluLeuProIlytyrProIlyValGlnAsnAsn	315
Qy	242	AAAGTCACATTTAAAGATGTCAGAAATAATGATTGATCTGAAGAGGATCATAGTGTAT	301
Db	316	ThrLeu	332
Qy	302	GAAGTC	307
Db	333	ThrAlaProProAsnGluLeuSerAlaAlaAsnGlyAspTyrTyrValLeuThrAsnLys	352
Qy	308	-----AGAAGCGTGGTATATTTGAA	328
Db	353	TrpIleThrMetProArgAsnGlySerAspLeuGlyGluGluArgLeuValThrValTyr	372
Qy	329	GGCACCAAGACTCCCTTTAGAGCATCACTGTACGTA	367
Db	373	AlaAsnArgAsp	389
Qy	367	-----	367
Db	390	CysValMetAlaLeuSerPheGlnPheSerIleAspAsnArgGlnLeuTyrValSerPro	409
Qy	368	-----GTCAGTTACGTAATACTCGAGAGGTGACAAGGCTGACT	406
Db	410	LysAspValArgGlyValPheProThrGluThrGlyPheLeuThrValLeuProHisLys	429
Qy	407	---GACCGTGGC---TACTCACAT	424
Db	430	HisAspAspGlyAsnIleTyrAsnHisValAlaHisValGluLeuAspGlyThrGlyThr	449
Qy	425	-----TCTTGTGTCATAGTCAGCAGCATGTGACTTCTTTATA	460
Db	450	GlyLysIleThrLysTrpIleGlyGluAsnPheAspValIleLeuValLeuGlyTyrSer	469
Qy	460	-----	460
Db	470	SerLysIleAspAlaLeuThrPheSerAlaTyrGlyAspGlyValGlyValGluPheSerThr	489
Qy	461	-----AGTAAGTATAGTAACAGAGAGATCCACACTGTGTGTCCCTTAC	505
Db	490	TyrIleValArgGluAlaMetTyrSerAsnLysLys	505
Qy	506	AAG-----CTATCAAGTCTCGAAGATGACCCAACT	535
Db	506	LysValThrAspGlnPheGluAspCysLysThrLeuGlySerGlnSerAlaAspProThr	525
Qy	536	-----TCCAAA	559
Db	526	GlyGlnArgIleValValGlnCysGluLysProPheAspAsnThrArgLeuTyrLeuVal	545
Qy	560	ACCATTTTGATTCCGCA	589
Db	546	AspValValAspThrThrLysLysIleMetLeuGluGlyGlyThrLysAlaValIlePro	565
Qy	590	GACTATACTCTCCAGAAATTTTC	646
Db	566	PheAspValProAsnMetLysPheGlyLysLeuLysLeuProSerGlyIleAspGlyHis	595
Qy	647	GGGATGCTCTACAGCCTCATGATCTACAGCCTGGAAGAAATATCTTACTGTCTGTTTC	706
Db	586	TyrMetMetLeuThrProAlaAsnLeuLeuAspGlyAlaLysIleProLeuLeuLeuAsp	605
Qy	707	ATATATGGTGTCTCAGGTGCGAGTTGGTGAATAATCGGTTTAAAGGAGTCAAGTATTC	766
Db	606	IleTyrGlyGlyProAspSerLysGlnValPheGlnLysThrProThrAlaHisAlaIle	625
Qy	767	CGCTTGAATACCCCTAGCCCTCTAGGTTATGTGGTTGTAGTATAGACACAGGGGATCC	826

Db 626 GlnIleValSerGlnTyTrpAspIleAlaTyR-----AlaArgIleAspValArgGlyThr 643
 QY 827 TGTCCACCGAGGGCTTAAATTTTGAAGGCGCTTTTAAATATATAAAATGGGTCAATAGAAAT 886
 Db 644 GlyGlyArgGlyTrpAspValLysGluAlaValTyArgLysLeuGlyAspAlaGluVal 663
 QY 887 GAGCATCAGGTGGAAGGAGTCCCATATCTAGCTTCTCGATATGATTTTCATTGACATTAGT 946
 Db 664 ValAspThrLeuAspMetIleArgAlaPheIleAsnThrPheGlyPheIleAspGluAsp 683
 QY 947 CGTGTGGGATCCACCGCTGCTCTATCGAGGATACCTC---TCCCTGATGGCATTAATG 1003
 Db 684 ArgIleAlaValMetGlyTrpSerTyGlyGlyPheLeuThrSerLysIleAlaIleLys 703
 QY 1004 CAGAGTTCAGATATCTTCAGGGTTGCTATTGCTGGGCGCCAGTCAGTCTGTGGGATCTTC 1063
 Db 704 AspGlnGlyGluLeuValLysCysAlaIleSerIleAlaProValThrAspPheLysTyR 723
 QY 1064 TATGATACAGGATACACG 1081
 Db 724 TyrAspSerAlaTyTrp 729

RESULT 13
 E87495
 prolyl oligopeptidase family protein [imported] - Caulobacter crescentus
 C:Species: Caulobacter crescentus
 C:Date: 20-Apr-2001 #sequence revision 20-Apr-2001 #text_change 09-Jul-2004
 C:Accession: E87495
 R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heide
 B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.
 n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fra
 Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
 A:Title: Complete Genome Sequence of Caulobacter crescentus.
 A:Reference number: AB7249; MUID:21173698; PMID:11259647
 A:Accession: E87495
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-683 <STO>
 C:Cross-references: UNIPROT:Q9A6U7; UNIPARC:UPI00000C7585; GB:AE005673; NID:gl3
 C:Genetics:
 A:Gene: CCI986

Alignment Scores:		
Pred. No.:	2,47e-12	Length: 683
Score:	226.50	Matches: 66
Percent Similarity:	45.9%	Conservative: 28
Best Local Similarity:	32.2%	Mismatches: 94
Query Match:	11.7%	Indels: 17
DB:	2	Gaps: 6

US-10-825-632-8 (1-1083) x E87495 (1-683)
 QY 377 GTAATCCTCGAGAGTGCAGAGGTGACCGTGGTACTCACATTCCTGTCGATC 436
 Db 356 ValLysThrGlyLysValSerAlaLeuThrGlyGluGlyHisValThrAlaPheAspVal 375
 QY 437 AGTCAGACATGTGACCTCTTTTATAAGTATAGTAAACAGAGAAATCCACACTGTGTG 496
 Db 376 GlyProSerGlyIleValPheAlaSer-----AspSerLeuLysSerProSer----- 391
 QY 497 TCCCTTTTACAAGCTATCAAGTCTCAAGATGACCCCACTTTGCAAAACAAAGGAATTTTGG 556
 Db 392 GluLeuPheLeuLeu-----ProAlaLysGlyProAlaValLysVal----- 405
 QY 557 GCCACCATTTTGGATTCAGCAGGTCTCTTCTGACTATATCTCTCCAGAAATTTTCTCT 616
 Db 406 AlaSerValSerSerGluAlaLeuLysAspValAlaTrpGlyGluProGluInPheSer 425
 QY 617 TTGAAAGTACTACTGGATTTACATTTGATGGGATGCTCTACAGCCCTCATGATCTACAG 676
 Db 426 PheLysGlyTrpAsnAspGluThrValHisGlyPheLeuValLysProAlaAsnPheAsp 445
 QY 677 CTGGAAGAATAATCCTACTGTGCTGTTCATATATATGTTGGTCTCTCAGGTGCAGTTGGTG 776

```
Db 446 ProAlaLysLysIleValProValAlaPheLeuLysGlyGlyProGln----- 461
Qy 737 AATAATCGGTTTAAAGGAGCAAGTATTTCCTGTAAT-----ACCTAGCCTCTCTA 790
Db 462 --GlySerPheSerAenAlaTrpSerTyraGtrPasnProGlnValTyraAenAla 480
Qy 791 GGTATCTGTTCTAGTGTAGACACACAGGAGCTCTGCACCGAGGCTTAAATTTGAA 850
Db 481 GlyTyraAlaValMetIleAspPheHisGlySerThrGlyTyraGlnAlaPheThr 500
Qy 851 GGCGCCTTTAAATATAAATGAGTCAATAGAAATGACATCAGGTGGAGGACTCCAA 910
Db 501 AspSerIleSerArgHisTrpGlyAspArgProLeuGluAspLeuGlnLysGlyTrpSer 520
Qy 911 TATCTAGCTTCTCGATATGATTTCATGACTTGTAGATCGTGGGCATCCACGCTGTTC 970
Db 521 PheValLeuSerLysTyraGlyPheLeuAspGlyAspArgAlaCysAlaLeuGlyAlaSer 540
Qy 971 TATGGAGGATACCTC 985
Db 541 TyrGlyGlyTyrMet 545
```

RESULT 14

A41793
dipeptidyl aminopeptidase-like protein (EC 3.4.14.-) short form, DPPX-S - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C:Accession: A41793

R:Wada, K.; Yokotani, N.; Hunter, C.; Doi, K.; Wenthhold, R.J.; Shimasaki, S.
Proc. Natl. Acad. Sci. U.S.A. 89, 197-201, 1992

A:Title: Differential expression of two distinct forms of mRNA encoding members of a dip
A:Reference number: A41793; PMID:92108018; PMID:1729689

A:Accession: A41793
A:Status: preliminary; not compared with conceptual translation

A:Molecule type: mRNA
A:Residues: 1-803 <WAD>

A:Cross-references: UNIPROT:P42659; UNIPARC:UPI000002A83C; GB:M76429; NID:g408719; PIDN:
A:Note: sequence extracted from NCBI backbone (NCBI:P:751138)

C:Superfamily: dipeptidyl-peptidase IV
C:Keywords: dipeptidylpeptide hydrolase; glycoprotein

F:257,342/Binding site: carbohydrate (Asn) (covalent) #status predicted

Alignment Scores:

Pred. No.:	7,95e-12	Length:	803
Score:	221.00	Matches:	103
Percent Similarity:	32.9%	Conservative:	58
Best Local Similarity:	21.0%	Mismatches:	169
Query Match:	11.4%	Indels:	160
DB:	2	Gaps:	16

US-10-825-632-8 (1-1083) x A41793 (1-803)

```
Qy 23 GGAGTCGCTACCTTGTCTCCAAGAGAA---TTTGATAGATATTCTGGCTATTGGTGG 79
Db 225 GlyLeuSerAspTrpLeuTyraGluGluLeuLysThrHisIleAlaHisTrp 244
Qy 80 TGTCCTCAAAAGCTGAACAACTCCAGTGGTGAATTTTGAATTTCTATATGAAGAA 139
Db 245 SerProAspGly-----ThrArgLeuAlaTyraAlaThr 255
Qy 140 AATGATGAATCTGAGTGGAAATTTTCATGTT-----ACATCCCTATGTTGAA 190
Db 256 IleAenAspSerArgValProValMetGluLeuProThrTyraGlySerValTyraPro 275
Qy 191 ACAAGGAGGAGATTCATTCGTTATCTTAAACAGGATACAGCAATCTTAAAGTCACT 250
Db 276 Thr-----AlaLysProTyraHisTyraProLysAlaGlyCysGluAsnProSerIleSer 293
Qy 251 TTT----- 253
Db 294 LeuHisValIleGlyLeuAenGlyProThrHisAspLeuGluMetThrProProAspAsp 313
```

```
Qy 254 ---AAGATGTCTAGAAATAATGATT----- 274
Db 314 ProArgMetArgGluTyraTyraIleThrMetValLysTrpAlaThrSerThrLysValAla 333
Qy 275 -----GATGCT----- 280
Db 334 ValAsnTrpLeuSerArgAlaGlnAenValSerIleLeuThrLeuCysAspAlaThrThr 353
Qy 280 ----- 280
Db 354 GlyValCysThrLysLysHisGluAspGluSerGluAlaTrpLeuHisArgGlnAenGlu 373
Qy 281 -----GAAGAAGG----- 289
Db 374 GluProValPheSerLysAspGlyArgLysPhePhePheValArgAlaIleProGlnGly 393
Qy 289 ----- 289
Db 394 GlyGlnGlyLysPheTyraHisIleThrValSerSerSerGlnProAenSerSerAsnAsp 413
Qy 290 -----ATCATAGTTGATGAA 304
Db 414 AsnIleGlnSerIleThrSerGlyAspTrpAspValThrLysIleLeuSerTyraAspGlu 433
Qy 305 GTCAGAGAGCTGGTATATTTGAAGGCACCAAGATCCCTTTAGAGCATCACCTGTAC 364
Db 434 LysArgSerGlnIleTyraPheLeuSerThrGluAspLeuProArgArgGlnLeuTyra 453
Qy 365 GTAGTCAGTTAGCTAAATCCTGGAGGTGACAGG-----CTGACT 406
Db 454 SerAlaSerThrVal-----GlySerPheAsnArgGlnCysLeuSerCysAspLeuVal 471
Qy 407 GACCGTGGCTACTCACCATTCTGTGTCATCAGTCAGCATCTGTGACTTCTTTATAAGTAAG 466
Db 472 AspAsnCysThrTyraPheSerAlaSerPheSerProGlyAlaAspPhePheLeuLeuLys 491
Qy 467 TATAGTAACAGAGAATCCACACTGTGTGCTTACAGCTATCAAGCTACCTGTGAGAT 526
Db 492 CysGluGlyProGlyValPro---ThrValSerValHisAenThrThrAspLysLysLys 510
Qy 527 GACCCCAACTTGCAAAACAAAGAA---TTTTGGGCCACCATTTTGGATTCAGCAGGTCCT 583
Db 511 MetPheAspLeuGluThrAsnGluHisValGlnLysAlaIleSerAspArgGlnMetPro 530
Qy 584 CTTCTCTGACTATACCTCCTCCAGAAATTTCTCTTTTGAAGAGTACTACTGGATTACATGG 643
Db 531 LysValGluTyraArgLysIleGlu-----ThrAspAspTyraAsnLeu 544
Qy 644 TATGGATGCTCTACAGCCTCATGATCTACAGCCTGCAAAAGAAATATCTACTGTGCTG 703
Db 545 ProIleGlnIleLeuLysProAlaThrPheThrAspThrAlaHisTyraProLeuLeuLeu 564
Qy 704 TTCATATATGTTGGTCTCAGGTGCTGTAATATCGGTTTAAAGAGTCAAGTAT 763
Db 565 ValValAspGlyThrProGlySerGlnSerValAlaGluLysPhe-----AlaValThr 582
Qy 764 TTCGCTTGAATACCTCTAGTCTCTAGTGTATGTGTGTGTGTGTGTGTGTGTGTGTGTGT 823
Db 583 TrpGluThrValMetValSerSerHisGlyAlaValValValLysCysAspGlyArgGly 602
Qy 824 TCTCTCACCGAGGCTTAAATTTGAAGGCGCTTTAAATATATAATATAATATATAATATAAT 883
Db 603 SerGlyPheGlnGlyThrArgLeuLeuHisGluValArgArgArgLeuGlySerLeuGlu 622
Qy 884 ATTGACGATCAGGTGGAGGAGTCCCAATATCTAGCTTCTCGATATGATTTCATTGATTA 943
Db 623 GluLysAspGlnMetGluAlaValArgValMetLeu---LysGluProTyraIleAspLys 641
Qy 944 GATCTGTGGGCTCCACGCTGCTCTATGAGGATACCTCTCTCTCTGATGAGCATTAATG 1003
Db 642 ThrArgValAlaValPheGlyLysAspTyraGlyTyraLeuSerThrTyraLeuLeuPro 661
Qy 1004 CAGAGGTCTAGAT-----ATCTTCAGGTTGCTATTGCTGGGCGCCCGAGTCACT 1051
```

```
Db      662 AlaLysGlyAspGlyGlnAlaProValPheSerCysGlySerAlaLeuSerProIleThr 681
QY      1052 CTGTGATCTTCTATCATACAGGATACAGC 1081
Db      682 AspPheLysLeuTyrAlaSerAlaPheSer 691

RESULT 15
B82580
alanyl dipeptidyl peptidase XP2260 [imported] - Xylella fastidiosa (strain 9a5c)
C;Species: Xylella fastidiosa
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C;Accession: B82580
R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A;Reference number: A82515; PMID:20365717; PMID:10910347
A;Note: for a complete list of authors see reference number A59328 below
A;Accession: B82580
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-709 <SIM>
A;Cross-references: UNIPROT:Q9PB84; UNIPARC:UPI00000C29B1; GB:AE004038; GB:AE003849; NID
A;Experimental source: strain 9a5c
R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H
as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigh
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XP2260

Alignment Scores:
Pred. No.: 9,66e-12 Length: 709
Score: 220.00 Matches: 73
Percent Similarity: 41.0% Conservative: 32
Best Local Similarity: 28.5% Mismatches: 99
Query Match: 11.4% Indels: 52
DB: 2 Gaps: 9

US-10-825-632-8 (1-1083) x B82580 (1-709)
QY      263 GAATATGATGATGCTGAGGAAGGATCATAGTTGATGAAGTCAGAGCGCTGGTATAT 322
Db      358 GluIleAlaLeuSerAlaAspGlyLysAlaLeu-----Tyr 369
QY      323 TTTCAGGACCAAGAACTCCCTTTAGAGCATCACCTGACGTAGTCAGTTACGTAAT 382
Db      370 ValAsnAla-----AspAspHisGlyGluHisProLeuPheLysValAsp---IleAla 386
QY      383 CTGAGAGGTGACAAAGGTGACTGACCGGTGCTACTCATCTTCTGCTGATCAGTCAG 442
Db      387 SerGlyLysValGluLysTrpValGlyGluGlySerValHisAlaProValLeuAlaGly 406
QY      443 CACTGTGACTCTTTATAGTAAGTATAGTAACACGAGAGATCCACACTGTGTGCTCCT 502
Db      407 -----GlyLysLeuAlaPheAlaArgAsn----- 414
QY      503 TACAAGCTATCAAGTCTCTGAAGATGACCCAACTTGCAAAACAAAGAAATTTGGGCCACC 562
Db      415 ---SerLeuLysSerAlaAsp-----GlnIlePheValThr 425
QY      563 ATTTGGATTCAGCAGGTCTCTTCTGCTGACTATCTCTCTCCA----- 604
```

Search completed: May 2, 2006, 01:38:29
Job time : 58.7491 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: May 2, 2006, 00:42:23 ; Search time 49.8855 Seconds
(without alignments)
3063.366 Million cell updates/sec

Title: us-10-825-632-8

Perfect score: 1938

Sequence: 1 ggaagaagatgccatcag.....tatgatacagatacacgga 1083

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 4332886

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp
-Q=/abss/ABSSWEB.spool/US10825632/runat_01052006_105946_3226/app_query.fasta_1
-DB=UniProt -QFMT=faстан -SUFFIX=n2p.rup -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abss02p
-USER=US10825632_@CGN_1_1_694_@runat_01052006_105946_3226 -NCPU=6 -ICPU=3
-NO MAP -NEG SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : UniProt_05.80.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	1836.5	94.8	898	1 DPP8 HUMAN	O6v1x1 homo sapien
2	1756.5	90.6	892	1 DPP8 MOUSE	Q80ya7 mus musculu
3	1220.5	63.0	863	1 DPP9 HUMAN	Q86ti2 homo sapien
4	1204.5	62.2	862	1 DPP9 MOUSE	Q8bvg4 mus musculu
5	1183.5	61.1	923	2 Q4SBM6 TETNG	Q4sbm6 tetraodon n
6	1182.5	61.0	847	2 Q6GR22 XENLA	O6gr22 xenopus lae
7	944	48.7	508	2 Q75273 HUMAN	O75273 homo sapien
8	718.5	37.1	1053	2 Q9VC20 DROME	Q9vc20 drosophila
9	718.5	37.1	1113	2 Q9VC19 DROME	Q9vc19 drosophila
10	705	36.4	886	2 Q7QBK1 ANOGA	Q7qbkl anopheles g
11	684	35.3	740	2 Q5TTK8 ANOGA	O5ttk8 anopheles g
12	630	32.5	621	2 Q7PTT8 ANOGA	O7ptt8 anopheles g
13	605.5	31.2	803	2 Q54U01 DIODI	Q54u01 dictyosteli
14	601.5	31.0	557	2 Q5TXJ2 ANOGA	O5txj2 anopheles g
15	527.5	27.2	738	2 Q9A6E0 CAUCR	Q9a6e0 caulobacter
16	494	25.5	763	2 Q8EAB7 SHEON	Q8eab7 shewanella

17	491.5	25.4	927	2	Q955K3_CABEL	Q955k3 caenorhabdi
18	491.5	25.4	931	2	O44987_CABEL	O44987 caenorhabdi
19	474	24.5	745	2	O6F317_9P8ED	O6f317 pseudomonas
20	471	24.3	751	2	O4UPD3_XANCP	O4upd3 xanthomonas
21	471	24.3	751	2	Q8P3V8_XANCP	Q8p3v8 xanthomonas
22	468	24.1	745	2	O5H5W8_XANOR	O5h5w8 xanthomonas
23	465.5	24.0	895	2	O6ICU7_CAEBR	O6icu7 caenorhabdi
24	463	23.9	743	2	O5QX36_IDILO	O5qx36 idiomarina
25	463	23.9	757	2	O8PFD7_XANAC	O8pfd7 xanthomonas
26	454	23.4	749	2	Q7NEK8_GLOVI	Q7nek8 gloeobacter
27	448	23.1	766	2	O4TNP1_9SPHN	O4tnp1 erythrobact
28	440	22.7	596	2	O6K880_ORYSA	O6k880 oryza sativ
29	430.5	22.2	741	2	P95782_XANNA	P95782 xanthomonas
30	422.5	21.8	746	2	O9FNF6_ARATH	O9fnf6 arabidopsis
31	418.5	21.6	778	2	O5IZP7_MAGGR	O5izp7 magnaporthe
32	409.5	21.1	739	2	O5FOY6_GLUOX	O5fgy6 gluconobact
33	403.5	20.8	720	2	O5NMW8_ZYMMO	O5nmw8 zymomonas m
34	403.5	20.8	732	2	Q7MUM6_PORGI	Q7muw6 porphyron
35	403	20.8	736	2	Q8A028_BACTN	Q8a028 bacteroides
36	390.5	20.1	552	2	O8GUJ7_ARATH	O8guj7 arabidopsis
37	383	19.8	719	2	O5LJ01_BACFN	O5lj01 bacteroides
38	382	19.7	719	2	Q65J00_BACFR	Q65j00 bacteroides
39	372	19.2	736	2	O5LGU5_BACFN	O5lgu5 bacteroides
40	372	19.2	736	2	O64XP9_BACFR	O64xp9 bacteroides
41	370.5	19.1	732	2	Q8A2Q1_BACTN	Q8a2q1 bacteroides
42	369	19.0	711	2	Q47900_FLAME	Q47900 flavobacter
43	366.5	18.9	730	2	Q93JY4_9BACT	Q93jy4 prevotella
44	365	18.8	809	2	Q4NVV1_9DBLT	Q4nvv1 anaeromyxob
45	353.5	18.2	731	2	Q6L872_PREIN	Q6l872 prevotella

ALIGNMENTS

RESULT 1

ID	DPP8 HUMAN	STANDARD;	PFT: 898 AA.
AC	O6V1X1; Q7Z4C8; Q7Z4D3; Q7Z4B1; Q8LWG7; Q8NEM5; Q96JX1; Q9HEM2;		
AC	Q9HBM3; Q9HEM4; Q9HBM5; Q9NXP4;		
DT	13-SEP-2005 (Rel. 48, Created)		
DT	13-SEP-2005 (Rel. 48, Last sequence update)		
DT	13-SEP-2005 (Rel. 48, Last annotation update)		
DE	Dipeptidyl peptidase 8 (EC 3.4.14.5) (Dipeptidyl peptidase VIII) (DP8)		
DE	(Prolyl dipeptidase DPP8) (Dipeptidyl peptidase IV-related protein 1)		
DE	(DPP-1).		
GN	Name=DPP8; Synonyms=MSTP097, MSTP135, MSTP141;		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;		
OC	Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 3), NUCLEOTIDE SEQUENCE [MRNA] OF		
RP	334-898 (ISOFORM 4), NUCLEOTIDE SEQUENCE [MRNA] OF 540-898 (ISOFORM		
RP	5), NUCLEOTIDE SEQUENCE [MRNA] OF 260-792 (ISOFORM 6), FUNCTION.		
RP	CATALYTIC ACTIVITY, ENZYME REGULATION, TISSUE SPECIFICITY, INDUCTION,		
RP	AND SUBCELLULAR LOCATION.		
RC	TISSUE=Placenta;		
RX	MEDLINE=20467194; PubMed=11012666;		
RA	Abbott C.A., Yu D.M.T., Woollatt E., Sutherland G.R., McCaughan G.W.,		
RA	Correll M.D.;		
RT	"Cloning, expression and chromosomal localization of a novel human		
RT	dipeptidyl peptidase (DPP) IV homolog, DPP8.";		
RL	Eur. J. Biochem. 267:6140-6150(2000).		
RN	[2]		
RP	NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 3), CATALYTIC ACTIVITY, ENZYME		
RP	REGULATION, BIOPHYSICO-CHEMICAL PROPERTIES, SUBCELLULAR LOCATION, AND		
RP	TISSUE SPECIFICITY.		
RC	TISSUE=Testis;...		
RX	PubMed=12662155; DOI=10.1042/BJ20021914;		
RA	Qi S.Y., Riviere P.J., Trojnar J., Junien J.-L., Akinaanya K.O.;		
RT	"Cloning and characterization of dipeptidyl peptidase 10, a new member		
RT	of an emerging subgroup of serine proteases.";		
RL	Biochem. J. 373:179-189(2003).		

CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 CC Muridae; Murinae; Mus.
 CC NCBI_TaxID=10090;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
 RC STRAIN=C57BL/6J; TISSUE=Testis;
 RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
 RA Okazaki Y., Furuno M., Kakuwaka T., Adachi J., Bono H., Kondo S.,
 RA Nikaide I., Otsu N., Saito R., Suzuki H., Yamanaka Y., Kiyosawa H.,
 RA Yagi K., Tomaru Y., Hasegawa Y., Nozaki A., Schombach C., Gojohori T.,
 RA Baldairelli R., Hilli D.P., Bult C., Hume D.A., Quakenbush J.,
 RA Schriber L.M., Kanapin A., Matsuda H., Bacalov S., Beisel K.W.,
 RA Blake J.A., Bradt D., Bruscia V., Chothia C., Corbani L.E., Cousins S.,
 RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
 RA Gasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
 RA Grimmond S., Guetlich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
 RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
 RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
 RA Megloun D.R., Maitais L., Marchionni L., McKenzie L., Miki H.,
 RA Negashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramchandran S.,
 RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
 RA Sandelin A., Schneider C., Sempile C.A., Setou M., Shimada K.,
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
 RA Verardo R., Wagner L., Walestedt C., Wang Y., Watanabe Y., Wells C.,
 RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang L., Yang L.,
 RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
 RA Hirozane-Kishikawa T., Kono H., Nakamura M., Sakazume N., Sato K.,
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
 RA Hara A., Haashizume W., Iotani K., Iehii Y., Itoh M., Kagawa I.,
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
 RA Yanunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
 RA Birney E., Hayashizaki Y.;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs";
 RL Nature 420:563-573 (2002).
 RN [2]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
 RC STRAIN=C57BL/6; TISSUE=Brain;
 RX MEDLINE=22389257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Straubeberg R.L., Feingold E.A., Grouse L.H., Derse J.G., Schuler G.D.,
 RA Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buerow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Brownstein M., Jost Ustin T.B., Toshitsuki S., Casavant T.B., Schetz T.E.,
 RA Raha S.S., Locuelli N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Besak S., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley A.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [3]
 RP PHOSPHORYLATION SITE TYR-325.
 RX PubMed=14729942; DOI=10.1074/mcp.D300003-MCP200;
 RA Shu H., Chen S., Bi Q., Mumbly M., Brekken D.L.;
 RT "Identification of phosphoproteins and their phosphorylation sites in
 RT the WEHI-231 B lymphoma cell line";
 RL Mol. Cell. Proteomics 3:279-286 (2004).
 CC -1- FUNCTION: Dipeptidyl peptidase that cleaves off N-terminal
 CC dipeptides from proteins having a Pro or Ala residue at position
 CC 2. May play a role in T-cell activation and immune function (By
 CC similarity).
 CC -1- CATALYTIC ACTIVITY: Release of an N-terminal dipeptide, Xaa-Yaa-|-

CC Zaa-, from a polypeptide, preferentially when Yaa is Pro, provided
 CC Zaa is neither Pro nor hydroxyproline.
 CC -1- ENZYME REGULATION: Inhibited by zinc. Inhibited by the serine
 CC proteinase inhibitor 4-(2-aminoethyl)benzenesulphonyl fluoride
 CC (AEBSEF), and by di-isopropylfluorophosphate. Specifically inhibited
 CC by isoinoline derivatives (By similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: Belongs to the peptidase S9B family. DPPIV subfamily.
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC EMBL; AK016546; BAB30295.2; -; mRNA.
 CC EMBL; BC043124; AAH43124.1; -; mRNA.
 CC EMBL; BC059222; AAH59222.1; -; mRNA.
 CC MEROPS; S09.018; -.
 CC Ensembl; ENSMUSG0000032393; Mus musculus.
 CC MGI; MGI:1921638; Dpp8.
 CC InterPro; IPR001375; Peptidase_S9.
 CC InterPro; IPR002469; Peptidase_S9B.
 CC InterPro; IPR000379; Ser_estrs.
 CC Pfam; PF00930; DPPIV_N; 1.
 CC Pfam; PF00326; Peptidase_S9; 1.
 CC K01654; Peptidase; Hydrolyase; Phosphorylation; Protease; Serine protease.
 CC ACT_SITE 749 749 Charge relay system (By similarity).
 CC ACT_SITE 827 827 Charge relay system (By similarity).
 CC ACT_SITE 859 859 Charge relay system (By similarity).
 CC MOD_RES 325 325 Phosphotyrosine.
 CC CONFLICT 87 87 G -> R (in Ref. 1).
 CC SEQUENCE 892 AA; 102186 MW; 59081CD9792803ED CRC64;
 Alignment Scores:
 Pred No.: 2,840-147 Length: 892
 Scores: 1756.50 Matches: 343
 Percent Similarity: 65.7% Conservative: 7
 Percent Local Similarity: 64.4% Mismatches: 10
 Query Match: 90.6% Indels: 173
 DB: 1 Gaps: 1
 US-10-825-632-8 (1-1083) x DPP8_MOUSE (1-892)
 QY 2 GAGAGAGATGCCAGATCAGCTGAGTCGCTACCTTTGTTCTTCAAGAGAAATTTGTATAGA 61
 Db 254 GluGluAspProArgSerAlaGlyValAlaThrPheValLeuGlnGluGluPheAspArg 273
 QY 62 TATCTGGCTATTGGTGGTGTCCAAAGCTGAAACAACTCCAGTGGTGGTAAATTTCTT 121
 Db 274 TyrSerGlyTyrTrpTrpCysProGlnAlaGluArgThrProSerGlyGlyLysLeu 293
 QY 122 AGAATCTATATCAAGAAATGATGAATCTGAGGTGGAATATTTCATGTTTACATCCCT 181
 Db 294 ArgileLeuTyrGluGluAsnAspGluSerGluValGluilelleHisValThrSerPro 313
 QY 182 ATCTTGAAACAGAGAGGAGCAGATTCATTCGGTTATCTTAAACAGGTACAGCAATCCT 241
 Db 314 MetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThrAlaAspPro 333
 QY 242 AAAGTCACCTTTTAAAGATGTCAGAAATATGATGATGTCGTGAAGGAGGATCATAT 295
 Db 334 LysValThrPheLysMetSerGluilelleValAlaAspAlaGlyGlyLysLeuAspVal 353
 QY 295 ----- 295
 Db 354 IleAspLysGluLeuValGlnProPheGluilellePheGluGlyValGluTyrIleAla 373
 QY 295 ----- 295
 Db 374 ArgAlaGlyTyrThrProGluGlyLysHisAlaTrpSerIleLeuLeuAspArgSerGln 393
 QY 295 ----- 295

Db 394 ThrHisLeuGlnIleValLeuIleSerProGluLeuPheIleProValGluAspAla 413
QY 295 ----- 295
Db 414 MetAspArgGlnArgLeuIleGluSerValProAspSerValThrProLeuIleIleTyr 433
QY 295 ----- 295
Db 434 GluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPheProGlnThr 453
QY 295 ----- 295
Db 454 HisGluAspGluIleGluPheIlePheAlaSerGluCysLysThrGlyPheArgHisLeu 473
QY 295 ----- 295
Db 474 TyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGlyGlyLeuPro 493
QY 295 ----- 295
Db 494 AlaProSerAspPheLysCysProIleLysGluGluIleThrIleThrSerGlyGluTrp 513
QY 296 -----GTTGATGAAGTCAGAGGCTGGTATAT 322
Db 514 GluValLeuGlyArgHisGlySerAsnIleTrpValAspGluAlaArgLysLeuValTyr 533
QY 323 TTTGAAGGACCAAGACTCCCTTTAGAGCATCCTGTAGTACGTAGTACGTAAAT 382
Db 534 PheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValThrSerTyrAlaAsn 553
QY 383 CTTGGAGAGTGACAGGCTGACTGACCGTGGTACTCACATCTTGTGTCATCAGTCAG 442
Db 554 ProGlyGluValValArgLeuThrAspArgGlyTyrSerHisSerCysLysSerArg 573
QY 443 CACTGTGACTCTTTTATAAGTAAAGTATAGTAAACAGAGAAATCCACACTGTGTGTCCTT 502
Db 574 HisCysAspPheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeu 593
QY 503 TACAAGTATCAAGTCTGAAGATGACCCCACTCCAAACAAAGGAATTTGGGCCACC 562
Db 594 TyrLysLeuSerSerProGluAspAspProValHisLysThrLysGluPheTrpAlaThr 613
QY 563 ATTTTGGATTACAGGCTCTCTCTGACTATCTCTCCAGAAATTTCTCTTTTGA 622
Db 614 IleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIlePheSerPheGlu 633
QY 623 AGTACTACTGGATTTACATTTGATGGATGCTCTCAAGCCTCATGATCTACAGCCTGGA 682
Db 634 SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly 653
QY 683 AAGAAATATCTACTGTGCTGTTCATATATGTTGCTCTCAGGTGCGATGTTGGTAATAT 742
Db 654 LysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeuValAsnAsn 673
QY 743 CGGTTTAAAGAGTCAAGTATTCGCTTGATACCTGACTGCTCTAGGTTAGTGGTT 802
Db 674 ArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrValVal 693
QY 803 GTACTGATACACACAGGGATCTCTGTCACCGAGGCTTAAATTTGAAGGCGCTTTAAA 862
Db 694 ValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLys 713
QY 863 TATAAATGGGTCAAAATAGAAATTTGACGATCAGGTGGAAGGACTCAATATCTAGCTTCT 922
Db 714 TyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyrLeuAlaSer 733
QY 923 CGATATGATTTCAATAGCTTAGATCTGTGGGCATCCAGGCTGCTCTATGAGGATAC 982
Db 734 GlnTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrSerTyrGlyGlyTyr 753
QY 983 CTCTCCCTGATGCATTAATGACAGGTCAGATATCTTCAGGGTTGCTATTGCTGGGGCC 1042

Db 754 LeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAlaGlyAla 773
QY 1043 CCAGTCACTCTGTGGATCTTTATGATACAGATACAG 1081
Db 774 ProValThrLeuTrpIlePheTyrAspThrGlyTyrThr 786
RESULT 3
DPP9 HUMAN
ID DPP9 HUMAN STANDARD; Q6ZMT2; Q6ZNU5; Q8N2J7; Q8N3F5; Q8WDXD8;
AC Q8GTI2; Q6A137; Q6UALO; Q6ZMT2; Q6ZNU5; Q8N2J7; Q8N3F5; Q8WDXD8;
AC Q8NT8; Q8BVR3;
DT 13-SEP-2005 (Rel. 48, Created)
DT 13-SEP-2005 (Rel. 48, Last sequence update)
DE Dipterydyl peptidase 9 (EC 3.4.14.5) (Dipeptidyl peptidase IX) (DPP)
DE (Dipeptidyl peptidase-like protein 9) (DPLP9) (Dipeptidyl peptidase
DE IV-related protein 2) (DPRP-2).
GN Name=DPP9; Synonyms=DPRP2;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 1), AND TISSUE SPECIFICITY.
RX MEDLINE=22347328; PubMed=12459266; DOI=10.1016/S0378-1119(02)01059-4;
RA Olsen C., Wagtmann N.;
RT "Identification and characterization of human DPP9, a novel homologue
RT of dipeptidyl peptidase IV.";
RL Gene 299:185-193(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 1), CATALYTIC ACTIVITY,
RP BIOPHYSICOCHEMICAL PROPERTIES, ENZYME REGULATION, TISSUE SPECIFICITY,
RP AND SUBCELLULAR LOCATION.
RC TISSUE=Colon;
RX PubMed=12662155; DOI=10.1042/BJ200201914;
RA Qi S.Y., Riviere P.J., Trojnar J., Junien J.-L., Akineanya K.O.;
RT "Cloning and characterization of dipeptidyl peptidase 10, a new member
RT of an emerging subgroup of serine proteases.";
RL Biochem. J. 373:179-189(2003).
RN [3]
RP NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 1), PARTIAL NUCLEOTIDE SEQUENCE
RP [MRNA] (ISOFORM 2), CATALYTIC ACTIVITY, BIOPHYSICOCHEMICAL PROPERTIES,
RP TISSUE SPECIFICITY, AND SUBCELLULAR LOCATION.
RX PubMed=15245913; DOI=10.1016/j.bbexp.2004.03.010;
RA Ajami K., Abbott C.A., McCaughan G.W., Gorrell M.D.;
RT "Dipeptidyl peptidase 9 has two forms, a broad tissue distribution,
RT cytoplasmic localization and DPPIV-like peptidase activity.";
RL Biochim. Biophys. Acta 1679:18-28(2004).
RN [4]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 1).
RC TISSUE=Placenta, and Skin;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
RA Diatchenko L., Marusina K., Farmer A.F., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo A.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Heltan E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield V.S.N., Krzywinski M.I., Skalska U., Smalish D.E.,
RA Scherch A., Schein J.E., Jones S.J.-M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

Db	254	ThrGlyTyrTrpTrpCysProThrAlaSerTrpGluGlySerGluGlyLeuLysThrLeu	273	614	HisThrArgSerAspValArgLeuTyrGlyMetIleTyrLysProHisAlaLeuGlnPro	633
Qy	122	AGAAATTCATATGATGAAGAAATGATGAATCTGAGGTGGAATATTATTCATGTTACATCCCT	181	680	CGAAGAATAATCTACTGCTGCTTCATATATATGCTGCTCCTCAGGTGCAGTTGGTGAAT	739
Db	274	ArgIleLeuTyrGluGluValAspGluSerGluValGluValIleHisValProSerPro	293	634	GlyLysLysHisProThrValLeuPheValTyrGlyGlyProGlnValGlnLeuValAsn	653
Qy	182	ATGTTGGAACAAGAGAGGCGAGTATTCATTCGGTTATCCTTAACACAGGTACACCAATCCT	241	740	AATCGGTTTAAAGAGGCAAGTATTTCCTTGAATACCTACCTCTCTAGGTATATGTG	799
Db	294	AlaLeuGluGluArgLysThrAspSerTyrArgTyrProArgThrGlySerLysAsnPro	313	654	AsnSerPheLysGlyIleLysTyrLeuArgLeuAsnThrLeuAlaSerLeuGlyTyrAla	673
Qy	242	AAAGTCACCTTTTAAAGATCTGACAAATAATGATTGATGCTGAAGAGAGATCATA	295	800	GTGTAGTAGTATACACACAGGGGATCTGTCCACCGAGGCTTAATTTTGAAGCGCCTTT	859
Db	314	LysIleAlaLeuLysLeuAlaGluPheGlnThrAspSerGlnGlyLysIleValSerThr	333	674	ValValValIleAspGlyArgGlySerCysGlnArgGlyLeuArgPheGluGlyAlaLeu	693
Qy	295	-----	295	860	AAATATAAAATGGGTCAAATAGAAATTCAGCATCAGGTGGAAGGACTCCATATCTAGCT	919
Db	334	GlnGluLysGluLeuValGlnProPheSerSerLeuPheProLysValGluTyrIleAla	353	694	LysAsnGlnMetGlyGlnValGluIleGluAspGlnValGluGlyLeuGlnPheValAla	713
Qy	295	-----	295	920	TCTCGATATGATTTTCATTGACTTAGATCGTGTGGGCATCCACGGCTGGTCTTATGGAGGA	979
Db	354	ArgAlaGlyTrpThrArgAspGlyLysTyrAlaTrpAlaMetPheLeuAspArgProGln	373	714	GluLysTyrGlyPheIleAspLeuSerArgValAlaIleHisGlyTrpSerTyrGlyGly	733
Qy	295	-----	295	980	TACCTCTCCCTCATGCGCATTAATGCAGAGTTCAGATATCTTCAGGGTTGCTATTGCTGGG	1039
Db	374	GlnTrpLeuGlnLeuValLeuLeuProProAlaLeuPheIleProSerThrGluAsnGlu	393	734	PheLeuSerLeuMetGlyLeuIleHisLysProGlnValPheLysValAlaIleAlaGly	753
Qy	295	-----	295	1040	GCCCCAGTCACCTCTGTGGATCTTCTATGATACAGGATACAGC	1081
Db	394	GluGlnArgLeuAlaSerAlaArgAlaValProArgAsnValGlnProTyrValTyr	413	754	AlaProValThrValTrpMetAlaTyrAspThrGlyTyrThr	767
Qy	295	-----	295	RESULT 4		
Db	414	GluGluValThrAsnValTrpIleAsnValHisAspIlePheTyrProPheProGlnSer	433	DP99 MOUSE		
Qy	295	-----	295	ID DP99 MOUSE STANDARD; PRT; 862 AA.		
Db	434	GluGlyGluAspGluLeuCysPheLeuArgAlaAsnGluCysLysThrGlyPheCysHis	453	AC Q8BVG4: Q6KAM9; Q8BWT9;		
Qy	295	-----	295	DT 13-SEP-2005 (Rel. 48, Created)		
Db	454	LeuTyrLysValThrAlaValLeuLysSerGlnGlyTyrAspTrpSerGluProPheSer	473	DT 13-SEP-2005 (Rel. 48, Last sequence update)		
Qy	295	-----	295	DT 13-SEP-2005 (Rel. 48, Last annotation update)		
Db	474	ProGlyGluAspGluPheLysCysProIleLysGluGluLeuAlaLeuThrSerGlyGlu	493	DE Dipeptidyl peptidase 9 (EC 3.4.14.5) (Dipeptidyl peptidase IX) (DP9)		
Qy	296	-----	296	DN Name=dp9;		
Db	494	TrpGluValLeuAlaArgHisGlySerLysIleTrpValAsnGluGluThrLysLeuVal	513	OS Mus musculus (Mouse)		
Qy	320	TATTTGAAGGCACCAAGACTCCCTTTAGAGCATCCTGTACGTAGTACGTACGTTACGTA	379	OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
Db	514	TyrPheGluGlyThrLysAspThrProLeuGluHisHisLeuTyrValValSerTyrGlu	533	OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;		
Qy	380	AATCTGGAGGTGACAGGCTGACTGACCGTGGCTACTCACATCTTCTGTCGATCAGT	439	OC Muroidae; Muridae; Mus;		
Db	534	AlaAlaGlyGluLeuValArgLeuThrThrProGlyPheSerHisSerCysSerMetSer	553	OX NCBI_TaxID=10090;		
Qy	440	CAGCATGTCATCTTTTATAAGTATAGTAAACCAAGAAGATCCACACTGTGTGTCC	499	RN NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 1).		
Db	554	GlnAsnPheAspMetPheValSerHisTyrSerSerValSerThrProProCysValHis	573	RC STRAIN=C57BL/6J; TISSUE=Liver, and Olfactory bulb.		
Qy	500	CTTTACAGCTATCAAGTCTGAGATGACCCCACTTCCAAACAAGGAATTTTGGGCC	559	RX MEDLINE=22354683; PubMed=1246851; DOI=10.1038/nature01266;		
Db	574	ValTyrLysLeuSerGlyProAspAspProLeuHisLysGlnProArgPheTrpAla	593	RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,		
Qy	560	ACCATTTGGATTACAGAGTCTCTCTCTGACTACTCTCTCCAGAAATTTTCTCTTT	619	RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,		
Db	594	SerMetMetGluAlaAlaSerCysProProAspTyrValProProGluIlePheHisPhe	613	RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,		
Qy	620	GAAAGTACTACTGGATTTTACATTGTATGGGATGCTCTCAAGCCCTCATGATCTTACACCT	679	RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojibori T.,		
				RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,		
				RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,		
				RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,		
				RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,		
				RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,		
				RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,		
				RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,		
				RA Konagaya A., Kurochkin I.V., Lee Y., Lienhard B., Lyons P.A.,		
				RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,		
				RA Nagashima T., Numa K., Okido T., Pavan W.J., Perte G., Pesole G.,		
				RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,		
				RA Ravasi A., Schneider C., Sempole C.A., Setou M., Shimada K.,		
				RA Sadelin A., Schneider C., Sempole C.A., Setou M., Shimada K.,		
				RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,		
				RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,		
				RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,		
				RA Yuan Z., Zavalan M., Zhu Y., Zimmer A., Carrincci P., Hayatsu N.,		
				RA Hirozane-Kishikawa T., Kono H., Nakamura M., Sakazume N., Sato K.,		
				RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,		
				RA Hara A., Hashizume W., Inotani K., Ishii Y., Itoh M., Kagawa I.,		
				RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,		

RA Yaeunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayaishizaki Y.,
RA "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs";
RL Nature 420:563-573 (2002).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 2).
RC TISSUE=Thymus;
RX PubMed=15449545;
RX Okazaki N., Kikuno R., Ohara R., Inamoto S., Koseki H., Hiraoka S.,
RA Saga Y., Kitamura H., Nakagawa T., Nagase T., Ohara O., Koga H.;
RT "Prediction of the coding sequences of mouse homologues of FLJ genes:
RT the complete nucleotide sequences of 110 mouse FLJ-homologous cDNAs
RT identified by screening of terminal sequences of cDNA clones randomly
RT sampled from size-fractionated libraries.";
RL DNA Res. 11:127-135 (2004).
RN [3]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 1).
RC STRAIN=FBV/N; TISSUE=Salivary gland;
RX MEDLINE=22388257; PubMed=1247932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Pahey J., Helton E., Kettelman M., Madan A.C., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Radríguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalek U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16999-16993 (2002).
CC -!- FUNCTION: Dipeptidyl peptidase that cleaves off N-terminal
CC dipeptides from proteins having a Pro or Ala residue at position 2
CC (By similarity).
CC -!- CATALYTIC ACTIVITY: Release of an N-terminal dipeptide, Xaa-Yaa-|-
CC Zaa, from a polypeptide, preferentially when Yaa is Pro, provided
CC Zaa is neither Pro nor hydroxyproline.
CC -!- ENZYME REGULATION: Inhibited by the serine proteinase inhibitor 4-
CC (2-aminoethyl)benzenesulphonyl fluoride (AEBSF), and by di-
CC isopropyl fluorophosphate (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytosolic (By similarity).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=QBVG4-1; Sequence=Displayed;
CC Name=2;
CC IsoId=QBVG4-2; Sequence=VSP_013870, VSP_013871, VSP_013872;
CC -!- SIMILARITY: Belongs to the peptidase S9B family. DPPIV subfamily.
CC -!- CAUTION: Ref.2 sequence differs from that shown due to a
CC frameshift in position 745.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; AK050021; BAC34034.1; -; mRNA.
CC EMBL; AK078301; BAC37211.1; -; mRNA.
CC EMBL; AK131178; BAD21428.1; ALT_SEQ; mRNA.
CC EMBL; BC057631; AAH57631.1; -; mRNA.
CC MEROPS; S09.019; -;
CC Ensembl; ENSMUSG0000001229; Mus musculus.
CC MGI; MGI:2443967; Dpp9.

DR InterPro; IPR001375; Peptidase_S9.
DR InterPro; IPR002469; Peptidase_S9B.
DR InterPro; IPR000379; Ser_estr.
DR Pfam; PF00930; DPPIV_N; 1.
DR Pfam; PF00326; Peptidase_S9; 1.
KW Alternative splicing; Amino peptidase; Hydrolase; Protease;
KW Serine protease.
FT ACT_SITE 729 729 Charge relay system (By similarity).
FT ACT_SITE 807 807 Charge relay system (By similarity).
FT ACT_SITE 839 839 Charge relay system (By similarity).
FT VARSP_LIC 1 75 Missing (in isoform 2).
FT VARSP_LIC 748 787 /FTID=VSP_013870.
FT VARSP_LIC 748 787 VAIAGAPVTVMAYDTGVTERVMDVPENNOQGYEAGSVAL
FT -> PHEAESPSLSLPATTDPRMASSSSMWEAKGTPEASEG
FT OR (in isoform 2).
FT /FTID=VSP_013871.
FT /FTID=VSP_013872.
FT /FTID=VSP_013872.1; BAC37211).
FT CONFLICT 369 369 D -> Y (in Ref. 1; BAC37211).
FT CONFLICT 546 546 S -> F (in Ref. 2).
FT CONFLICT 777 777 O -> K (in Ref. 1; BAC37211).
SQ SEQUENCE 862 AA; 98001 MW; B1D565E824A834E8 CRC64;
Alignment Scores:
Pred. No.: 3 85e-98 Length: 862
Score: 1207.50 Matches: 234
Percent Similarity: 53.7% Conservatives: 53
Best Local Similarity: 43.8% Mismatches: 72
Query Match: 62.2% Indels: 175
DB: Gaps: 2
US-10-825-632-8 (1-1083) x DPP9_MOUSE (1-862)
QY 5 GAAGATGCCAGATCAGCTGGAGTCTGCTTCTTCCAGAGAAATTTGATAGATAT 64
D 233 ASPASNPROLYSERSALAGLYVALALATHRPHEVALLEGInGluGluPheAspArgpHe 252
QY 65 TCTGGCTATTGTGTGTCTCCAAAAGCTCCCACTCCCACTGGTGTGTGT---AAAATTCTT 121
D 253 ThrGlyCysTrpTrpCysProThrAlaSerTrpGluGlySerGluGlyLeuLeu 272
QY 122 AGAATTCTATATGAAGAAATGATGATCTGAGTGGAAATTTATTCATGTTTACATCCCT 181
D 273 ArgIleLeuTrpGluGluValAspGluSerGluValGluValIleHisValProSerPro 292
QY 182 ATGTTGGAAACAGAGGAGGAGGAGTTCCTTCCGTTATCTTAAACAGTACAGCAATCCT 241
D 293 AlaLeuGluGluValArgLysThrAspSerTrpArgTrpProArgThrGlySerLysAsnPro 312
QY 242 AAAGTCACCTTTTAAAGATGTCAGAAATATGATGATGCTCGAAGGAGGATCATA----- 295
D 313 LysIleAlaLeuLysLeuAlaGluLeuGlnThrAspHisGlnGlyLysIleValSerSer 332
QY 295 ----- 295
D 333 CysGluLysGluLeuValGlnProPheSerSerLeuPheProLysValGluTrpIleAla 352
QY 295 ----- 295
D 353 ArgAlaGlyTrpThrArgAspGlyLysTrpAlaTrpAlaMetPheLeuAspArgProGln 372
QY 295 ----- 295
D 373 GlnArgLeuGlnLeuValLeuLeuProProAlaLeuPheIleProAlaValGluSerGlu 392
QY 295 ----- 295
D 393 AlaGlnArgGlnAlaAlaAlaArgAlaValProLysAsnValGlnProPheValIleTrp 412
QY 295 ----- 295
D 413 GluGluValThrAsnValTrpIleAsnValHisAspIlePheHisPropheProGlnAla 432

QY 295 ----- 295
 Db 433 GluGlyGlnGlnAspPheCysPheLeuArgAlaAsnGluCysLysThrGlyPheCysHis 452
 QY 295 ----- 295
 Db 453 LeuTyrArgValThrValGluLeuLysThrLysAspTyrAspTrpThrGluProLeuSer 472
 QY 295 ----- 295
 Db 473 ProThrGluAspGluPheLysCysProLleLysGluGluValAlaLeuThrSerGlyGlu 492
 QY 296 ----- 319
 Db 493 TrpGluValLeuSerArgHisGlySerLysLleTrpValAsnGluGlnThrLysLeuVal 512
 QY 320 TATTGAGGACCAAGACTCCCTTTAGAGCATCACCTGTAGTGTAGTACAGTACGTTAGCTA 379
 Db 513 TyrPheGlnGlyThrLysAspThrProLeuGluHisLysLeuTyrValValSerTyrGlu 532
 QY 380 AATCCTGAGAGGTGACAGGCTGACTGACCGTGTCTACTCATCTTCTGTGTCATCAGT 439
 Db 533 SerAlaGlyGluLleValArgLeuThrThrLeuGlyPheSerHisSerCysSerMetSer 552
 QY 440 CAGCACTGTGACTCTTTTAAAGTATAGTATAGTAACCAAGAAATCCACACTGTGTGCC 499
 Db 553 GlnSerPheAspMetPheValSerHisTyrSerSerValSerThrProProCysValHis 572
 QY 500 CTTTACAGCTATCAAGTCTGAGATGACCCAACTTGCACAAACAAAGAAATTTGGGCC 559
 Db 573 ValTyrLysLeuSerGlyProAspAspProLeuHisLysGlnProArgPheTrpAla 592
 QY 560 ACCATTGTTGGATTGACAGGCTCTCTTCTGATATATCTACTCTCCACGAAATTTCTCTTT 619
 Db 593 SerMetMetGluAlaAlaAsnCysProAspTyrValProGluLlePheHisPhe 612
 QY 620 GAAATCTACTGATTTACATTTGATGATGATGATGATGATGATGATGATGATGATGATG 679
 Db 613 HisThrArgAlaAspValGlnLeuTyrGlyMetLleTyrLysProHisThrLeuGlnPro 632
 QY 680 GGAAGAATAATCTCTACTGCTGCTTTCATATATGTTGTTGTTGTTGTTGTTGTTGTTG 739
 Db 633 GlyArgLysHisProThrValLeuPheValTyrGlyGlyProGlnValGlnLeuValAsn 652
 QY 740 AATCGGTTTAAAGAGTCAAGTATTTCCGTTGATACCTACCTGCTCTGATGTTATG 799
 Db 653 AsnSerPheLysGlyLleLysTyrLeuArgLeuAsnThrLeuAlaSerLeuGlyTyrAla 672
 QY 800 GTTGTAGTATAGACAAACAGGGATCTCTGTCACGAGGCTTAAATTTGAAGCGCTTT 859
 Db 673 ValValValLleAspGlyArgGlySerCysGlnArgGlyLeuHisPheGluGlyAlaLeu 692
 QY 860 AATATATAAATGGTCAATAGAAATTCAGATCAGGTGGAAGGACTCCATATCTAGCT 919
 Db 693 LysAsnGlnMetGlyGlnValGluLleGluAspGlnValGluGlyLeuGlnTyrValAla 712
 QY 920 TCTGATATGATTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 979
 Db 713 GluLysTyrGlyPheLleAspLeuSerArgValAlaLleHisGlyTrpSerTyrGlyGly 732
 QY 980 TACTCTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1039
 Db 733 PheLeuSerLeuMetGlyLeuLleHisLysProGlnValPheLysValAlaLleAlaGly 752
 QY 1040 GCCCAGTCTCTGTTGATCTTCTATGATACAGTACAG 1081
 Db 753 AlaProValThrValTrpMetAlaTyrAspThrGlyTyrThr 766
 RESULT 5
 Q4SBM6_TETNG
 ID Q4SBM6_TETNG PRELIMINARY; PRT; 923 AA.
 AC Q4SBM6;
 DT 13-SEP-2005 (TrEMBLrel. 31, Created)

DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
 DE Chromosome 15 SCF14667, whole genome shotgun sequence.
 DE (Fragment).
 GN ORFNames=GSTENG0020903001;
 OS Tetraodon nigroviridis (green puffer).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
 OC Tetraodontoidea; Tetraodontidae; Tetraodon.
 OX NCBI_TaxID=99883;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Jallou O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
 RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
 RA Nicoud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
 RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
 RA Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
 RA Bionet C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
 RA Cruaud C., Duprat S., Brothier P., Coutanceau J.P., Gouzy J., Bosak S.,
 RA Parra G., Lardier J.N., Guigo R., Zody M.C., Mesirov J.,
 RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
 RA Wincker P., Lander E.S., Weissenbach J., Roest Crolius H.,
 RA "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
 RT the early vertebrate proto-karyotype."
 RL Nature 431.946-957(2004).
 [2]
 RP NUCLEOTIDE SEQUENCE.
 RG Genoscope; Whitehead Institute Centre for Genome Research;
 RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
 CC -!- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 DR EMBL; CA901014667; CAG01956.1; -; Genomic_DNA.
 FT NON_TER 1
 FT NON_TER 923
 SQ SEQUENCE 923 AA; 105211 MW; 139702382004009 CRC64;
 Alignment Scores:
 Pred. No.: 2,88e-96 Length: 923
 Score: 1183.50 Matches: 237
 Percent Similarity: 49.1% Conservative: 59
 Best Local Similarity: 39.3% Mismatches: 64
 Query Match: 61.1% Indels: 243
 DB: 2 Gaps: 2
 US-10-825-632-8 (1-1083) x Q4SBM6_TETNG (1-923)
 QY 2 GAAGAAGATCCAGATCAGCTGGAGTCCTACCTTTGTTCTCCAAAGAAATTTGATAGA 61
 Db 225 LysGluAspProLysSerAlaGlyValAlaThrPheValLleGlnGluPheAspArg 244
 QY 62 TATTCTGGCTATTGGTGGTGTCCAAAAGCTCAAAACAACTCCAGTGGTGTAAATCTTT 121
 Db 245 PheThrGlyTyrTrpSerProSerAlaValGluAspProAspGlyGlyLysArgVal 264
 QY 122 AGAATCTTATGAGAAATATGATGATCTGAGGTGGAATATTATTTACATCCCT 181
 Db 265 TyrLeuLeuTyrGluGluValAspGluThrGluValGluLleHisValProSerPro 284
 QY 182 ATGTGGAACACAGAGGAGGACAGATTCATTCCTGTTATCTCTAAACAGGTACAGCAATCT 241
 Db 285 AlaLeuGluGluArgLysAlaAspAlaTyrArgTyrProArgThrGlySerLysAsnPro 304
 QY 242 AAAGTCATCTTTTAAAGATGTGAGAAATATGATGATGTGCTGAAGGAGG 289
 Db 305 GlnAlaThrLleLysLeuValGluLleLysThrAspGlnGlnGlyArgValSerLeuCys 324
 QY 289 ----- 289
 Db 325 ArgLeuLeuCysValPheLeuLeuArgSerAspArgAlaValLeuThrCysLeuTyr 344

QY 289 ----- 289
 Db 345 LeuGlnIleValSerThrGlnAspLysGluLeuAlaValProPheThrSerLeuPhePro 364
 QY 289 ----- 289
 Db 365 GlyThrGluTyrIleAlaArgValGlyTrpThrSerAspGlyLysTyrGlyTrpAlaAla 384
 QY 289 ----- 289
 Db 385 LeuLeuAspArgSerGlnArgLysLeuGlnLeuValLeuLeuProAlaPhePheVal 404
 QY 289 ----- 289
 Db 405 ProValThrAspAspProAlaArgGlnGlnSerLeuGluAlaValProAspArgThr 424
 QY 289 ----- 289
 Db 425 GlnProTyrValIleTyrGluGluThrThrAspValTrpIleAsnValHisAspIlePhe 444
 QY 289 ----- 289
 Db 445 TyrProPheValGlnThrAlaGluAspGluPheThrPheIleTrpValAsnGluSerLys 464
 QY 289 ----- 289
 Db 465 ThrGlyPheSerHisLeuTyrLysIleThrSerValLeuHisProGlyPheHisCysTrp 484
 QY 289 ----- 289
 Db 485 AlaGluAlaTyrHisHisThrGluGlyAspProGlnArgArgIleProAlaValSerThr 504
 QY 289 ----- 289
 Db 505 AspValProProGlyAspPheLysCysAlaValLysGluGluIleThrLeuThrSerGly 524
 QY 290 -----ATCATAGTTGATGAGTCAAGTCAGAGGCTG 316
 Db 525 GluTrpGluValLeuAlaArgHisGlySerLysIleTrpValAsnGluSerSerLysLeu 544
 QY 317 GTATATTGTAAGCCACCAAGACTCCCTTTAGAGCATCACCTGCTAGCTAGTACAGTTAC 376
 Db 545 ValTyrPheGlnGlyThrArgAspThrProLeuGluHisLeuTyrValValSerTyr 564
 QY 377 GTAAATCCTGAGAGGTGACCAAGCTGACTGACCGTGGCTACTCATCTTCTGCTGATC 436
 Db 565 AspSerProGlyAspValValAlaGluLeuThrLysProGlyPheSerHisSerCysSerVal 584
 QY 437 AGT----- 439
 Db 585 SerGlnValLysLysSerLeuGlnSerAspTyrPhePheAsnTyrSerSerIleThrLeu 604
 QY 440 -----CAGCACTGTGACTTCTTTATAAGTAAGTAPAGT 472
 Db 605 ProLeuSerLeuSerSerPheIleTrpGlnAsnPheAspPhePheValSerHisTyrSer 624
 QY 473 AACCAAGAATCCACACTGTGTGCTCCCTTTACAGCTATCACTGCTGAAGATGACCCA 532
 Db 625 SerValCysThrProProCysValHisValTyrLysLeuAsnSerSerGluSerAspPro 644
 QY 533 ACTTGCAAAACAAAGAAATTTGGGCCACCAATTTGGATTACAGAGTCTCTCTCTGAC 592
 Db 645 LeuHisIleValProGluIleThrPheIleThrSerMetGluSerSerGlyCysProGlyAsp 664
 QY 593 TATACCTCTCCAGAAATTTCTTTTGAAGTACTACTGGAATTACATTGATGAGGAT 652
 Db 665 TyrSerProGluIlePheAspPheGlnGlyLysSerGlyPheGlnLeuTyrGlyMet 684
 QY 653 CTCACAGCCTCATGATCTACAGCTGGAAGAAATATCTACTGCTGTGTTTCATATAT 712
 Db 685 ValTyrLysProHisSerLeuGlnProGlyArgLysHisProThrValLeuPheValTyr 704

QY 713 GGTGCTCTCAGGTGAGTGGTGAATAATCGTTTAAAGGAGTCAAGTATTTCCGCTTG 772
 Db 705 GlyGlyProGlnValGlnLeuValAsnAsnSerPheLysGlyMetLysTyrLeuArgLeu 724
 QY 773 AATACCTCCTCCTCTAGGTATGTGGTTGTAGTAGATAGACAAACAGGGGATCTGTTCAC 832
 Db 725 AsnThrLeuAlaSerLeuGlyTyrAlaValValIleAspGlyArgGlySerCysGln 744
 QY 833 CGAGGGCTTAAATTTCAAGCGCTTTAAATAATAAATGCGTCAATAGAAATTCAGCAT 892
 Db 745 ArgGlyLeuGluPheGluSerAlaLeuLysAsnLysMetGlyGlnValGluIleGluAsp 764
 QY 893 CAGGTGGAAGGACTCCCAATATCTAGCTCTCGATATGATTTTCATTCAGTTCGTGTG 952
 Db 765 GlnValGluGlyLeuGlnTyrValAlaGluLysPheAsnPheValAspLeuSerArgVal 784
 QY 953 GGCATCCAGGCTGCTCCCTATGAGGATACCTCTCCCTGATGGCATTAATGACAGGTCA 1012
 Db 785 AlaIleHisGlyTyrSerTyrGlyGlyPheLeuSerLeuMetGlyLeuIleGlnArgPro 804
 QY 1013 GATATCTTCAGGTTGCTATTCTGCGGCCCACTGCTGTGGATCTTCTATGATACA 1072
 Db 805 AsnValPheLysLeuAlaIleAlaGlyAlaProValThrValTrpMetAlaTyrAspThr 824
 QY 1073 GGATACAGC 1081
 Db 825 GlyTyrThr 827
 RESULT 6
 Q6GR22 XENLA PRELIMINARY; PRT; 847 AA.
 ID Q6GR22 XENLA PRELIMINARY; PRT; 847 AA.
 AC Q6GR22;
 DT 05-JUL-2004 (TREMBLrel. 27, Created)
 DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
 DE MCB81313 protein.
 GN Name=MG81313;
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
 OC Xenopodinae; Xenopus; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Embryo;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.B.,
 RA Raha S.S., Loquellano N.A., Toehiyuki S., Carninci F., Frange C.,
 RA Bobak S.A., McEwan P.J., Peters G.J., Abramson K.D., Mullahy S.J.,
 RA Richards S., Worley K.C., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Blakesley R.W., Touchman J.W., Young A.C., Madan A., Rodriguez S., Sanchez A.,
 RA Buterfield Y.S.N., Grimwood J., Schmutz J., Myers R.M.,
 RA Roderich A.C., Touchman J.W., Jones S.J.M., Skalska U., Smailus D.E.,
 RA "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Embryo;
 RX MEDLINE=22394132; PubMed=12454917; DOI=10.1002/dvdy.10174;
 RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
 RA Richardson P.,
 RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus


```
RT      initiative";
RL      Dev. Dyn. 225:384-391(2002).
RN      [3]
RP      NUCLEOTIDE SEQUENCE.
RC      TISSUE=Embryo;
RA      Klein S., Gerhard D.S.;
RL      Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.
DR      EMBL; BC071112; AH71112.1; -, mRNA.
DR      GO; GO:0003824; F: catalytic activity; IEA.
DR      GO; GO:0008236; F: serine-type peptidase activity; IEA.
DR      GO; GO:0006508; F: proteolysis and peptidolysis; IEA.
DR      InterPro; IPR001375; Peptidase S9.
DR      InterPro; IPR002469; Peptidase S9B.
DR      InterPro; IPR000379; Ser_estr.
DR      Pfam; PF00930; DPPIV_N; 1.
DR      Pfam; PF00326; Peptidase S9; 1.
SQ      SEQUENCE 847 AA; 97484 MW; F515F2609B267BAF CRC64;

Alignment Scores:
Pred. No.:      3,48e-96      Length:      847
Score:          1182.50      Matches:    230
Percent Similarity: 54.0%      Conservative: 58
Best Local Similarity: 43.2%      Mismatches: 68
Query Match:      61.0%      Indels:    177
DB:              2          Gaps:      2

US-10-825-632-8 (1-1083) x Q6GR22_XENLA (1-847)
QY      5 GAAGATGCCAGATCAGCTCGAGTCCTACCTTTGTTCTCCAAAGAAGAAATTGTATAGATAT 64
DB      222 ASPASPProCysSerAlaGlyValAlaThrPheValIleGlnGluPheAspArgPhe 241
QY      65 TCTGGCTATTGGTGGTGTCCAAAGCTGAAACAACTCCAGCTGGTGTGTAATAAATCTTAGA 124
DB      242 ThrGlyTyrTrpTrpSerProThrValValAspLeuProGlyGlyPheSerValLeuArg 261
QY      125 ATCTATATGAGAAATATGAACTGAGGTGGAAATATTATGATGTTACATCCCTATG 184
DB      262 IleLeuTyrGluGluValAspGluSerAspValGluIleHisValProSerProAla 291
QY      185 TTGGAACAGGAGGACGATTCATTCCGTTATCTCTAAACAGGTACAGCAAAATCTCTAAA 244
DB      282 LeuGluGluArgGlyThrAspAlaTyrArgTyrProArgThrGlySerLysAsnProLys 301
QY      245 GTCACITTTTAAAGATGTGAGAAATATGATTGATGCTGAAAGAGGATCATA----- 295
DB      302 IleThrLeuLysLeuSerGluIleThrThrAsnAsnAspGlyGluIleValAlaAlaGln 321
QY      295 ----- 295
DB      322 AspLysGluLeuValLeuProPheHisSerLeuPheProSerAlaGluTyrIleValArg 341
QY      295 ----- 295
DB      342 AlaGlyTrpSerGlnAspGlyLysPheAlaTrpAlaMetLeuLeuAspArgProGlnCys 361
QY      295 ----- 295
DB      362 LeuLeuGlnLeuValLeuLeuProGluLeuPheIleProValCysGluAspGluGlu 381
QY      295 ----- 295
DB      382 GlnArgIleGlnTyrValLysLysValProGlnHisIleGlnProTyrValIleTyrGlu 401
QY      295 ----- 295
DB      402 GlnLysSerAspValTrpValAsnValHisAspValValHisLeuLeuArgGlnGluLys 421
QY      295 ----- 295
DB      422 GluAsnGluIleArgPheIleCysSerIleGluSerGlnThrGluPheCysHisLeuTyr 441
QY      295 ----- 295

442 LeuMetThrSerLeuLeuGlnAsnGluAsnArgGluTrpSerGlnMetThrIleProCys 461
295 ----- 295
462 ThrValAspMetPheLysCysGlnIleAsnGlnGluValAlaLeuThrSerGlyGluTrp 481
296 ----- 296
482 GluValLeuCysArgHisGlyAlaLysValTrpValAsnGluAlaMetGlnLeuValTyr 501
323 TTTGAAAGCACCACAAAGACTCCCTTTAGAGCATACCTGTAGTACGTAGTACGTAGTAAAT 382
502 PheGlnGlyThrLysAspThrProLeuGluHisHisLeuTyrValThrSerTyrGlnHis 521
383 CCTGGAGAGGTGACAGGCTGACTGACCGTGGCTACTCACATTTCTGCTCAGTCAGTCAG 442
522 ProGlyGluValValArgLeuThrGluArgGlyTyrSerHisSerCysThrMetSerPro 541
443 CACTGTGACTCTTTTATAAGTATAGTAAACAGAGAAGATCCACACTGTGTGTCCTTT 502
542 AsnPheAspMetPheValSerGlnTyrSerSerValSerCysProCysValHisLeu 561
503 TACAAGCTATCAAGTCTCTGAAGATGACCCCACTTCCAAACAAAGGAATTTTGGCCACC 562
562 TyrArgLeuAsnGly-----AsnProLeuTyrGlnHisProGlnPheTrpAlaSer 578
563 ATTTTGGATTTCAGCAGGCTCTCTCTGACTATATCTCTCCAGAAATTTTCTCTTTTGA 622
579 ValMetGluAlaAlaArgCysProProTyrTyrValProProGluIlePheHisPheGln 598
623 AGTACTACTGGATTTACATTGTATGGATGCTCTCAAGCCTCATGATCTACAGCTGGA 682
599 AlaSerSerGlnValLysLeuTyrGlyMetValTyrLysProHisAsnLeuValProGly 618
683 AAGAAATATCTACTGTGCTGTATATATGTTGTTCTCTCAGTCCAGTCCAGTTCGTGAATAT 742
619 ThrLysHisProThrValLeuPheValTyrGlyGlyProGlnValGlnLeuValAsnAsn 638
743 CGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCTAGCTCTCTAGGTTATGTGTT 802
639 SerPheLysGlyMetLysTyrLeuArgLeuAsnThrLeuAlaHisLeuGlyTyrAlaVal 658
803 GTAGTGATGACACAAAGGAGTCTCTGACAGGAGGCTTAAATTTTGAAGGCGCTTTAAA 862
659 ValValIleAspGlyArgGlySerCysHisArgGlyLeuAlaPheGluGlyAlaLeuLys 678
863 TATAAATGGGTCAAAATAGAAATTTGACATCAGGTGGAAGGACTCCAATATCTAGCTTCT 922
679 AsnLysMetGlyGlnValGluIleGlnAspGlnValGluGlyLeuHisPheValAlaGlu 698
923 CGATATGATTTTCAATTGATTCGTGTGGGCATCCAGCGTGTCTCTATGAGGATAC 982
699 ArgPheGlyPheValAspLeuAspArgValSerIleHisGlyTyrPsrSerTyrGlyPhe 718
983 CTCTCCCTGATGGCATTAAATGACAGAGTCAGATATCTTCAGGCTGTCTATTCGGGGCC 1042
719 LeuSerLeuMetGlyIleIleGlnArgProGluValPheLysValAlaIleAlaGlyAla 738
1043 CCAGTCACTCTGTGGATCTTCTATGATACAGGATACACG 1081
739 ProValThrLeuTrpMetAlaTyrAspThrGlyTyrThr 751

RESULT 7
O75273 HUMAN
ID O75273 HUMAN PRELIMINARY; PRT; 508 AA.
AC O75273;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE R26984.1 (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```


RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamis I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RA "The genome sequence of *Drosophila melanogaster*.";
RT Science 287:2185-2195(2000).
RL [2]
RN
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426065; PubMed=12537568;
RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA George R.A., Hoskins R.A., Laverty T., Muzny D.M., Nelson C.R.,
RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weinstein G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RA "Finishing a whole-genome shotgun: release 3 of the *Drosophila*
RT *melanogaster* euchromatic genome sequence.";
RT Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RL [3]
RN
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426070; PubMed=12537573;
RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J.W., Svirskas R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celniker S.E.;
RT "The transposable elements of the *Drosophila melanogaster* euchromatin:
RT a genomic perspective.";
RT Genome Biol. 3:RESEARCH0084.1-RESEARCH0084.20(2002).
RL [4]
RN
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a
RT systematic review.";
RT Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
RL [5]
RN
RP NUCLEOTIDE SEQUENCE.
RX Berkeley *Drosophila* Genome Project;
RA Celniker S., Carlson J., Wan K., Pfeiffer B., Frise E., George R.,
RA Hoskins R., Stapleton M., Pacleb J., Park S., Svirskas R., Smith E.,
RA Yu C., Rubin G.;
RT "Drosophila melanogaster release 4 sequence.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [6]
RP NUCLEOTIDE SEQUENCE.
RX FlyBase;
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
RN [7]
RN
RP NUCLEOTIDE SEQUENCE.
RX STRAIN=Berkeley;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dreenek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celniker S.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
CC -1- INTERACTION;
CC Q9VZ00.CG12016; NbExp=1; IntAct=EBI-105926, EBI-152819;
CC EMBL; AT0003749; AAF56357.2; -; Genomic_DNA.
DR EMBL; BT001499; AAN71254.1; -; mRNA.

DR IntAct; Q9VZ00; -;
DR MEROPS; S09.016; -;
DR Ensembl; CG3744; *Drosophila melanogaster*.
DR FlyBase; FBgn0039240; CG3744.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0004274; F:dipeptidyl-peptidase IV activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001375; Peptidase_S9.
DR InterPro; IPR002469; Peptidase_S9B.
DR InterPro; IPR000379; Ser_estrs_2.
DR Pfam; PF00930; DPPIV_N; 1.
DR Pfam; PF00326; Peptidase_S9; 1.
SQ SEQUENCE 1053 AA; 118053 MW; C94AA663AB464577 CRC64;

Alignment Scores:
Pred. No.: 7,21e-55 Length: 1053
Score: 718.50 Matches: 176
Percent Similarity: 41.1% Conservative: 55
Best Local Similarity: 31.3% Mismatches: 116
Query Match: 37.1% Indels: 215
DB: 2 Gaps: 10

US-10-825-632-8 (1-1083) x Q9VZ00_DROME (1-1053)

QY 5 GAAGATGCCAGATCAGCTGGAGTCGTACCTTTGTTCTCCAGAGAATTTGATGATAT 64
DB 409 AspAspAlaLeuSerAlaGlyValProSerTyrValMetGlnGluGluPheSerArgTyr 428
QY 65 TCTGCTATTGCTGCTGCCAAAAGCTGAAACAACTCCAGTGGTGGTAAATTCCTTACA 124
DB 429 GlnGlyPheTrpTrpGlnProHis-----SerAsnAspGlyIleTyrArg 443
QY 125 ATTCTATATGAAGAAATGATGATTCAGTGGAGTAAATTTATCATGTTACATCCCATG 184
DB 444 IleValTyrGluGluValAspGluSerGluValSerValTyrThrPheProSerSerThr 463
QY 185 TTGGAACACAGGAGGCGAGATTCATTCCTGTTATCTTAAACAGGTACACCAATCCTAAA 244
DB 464 AlaMetHisGlyArgValAspGluTyrArgPheProArgThrGlySerProAsnAlaLys 483
QY 245 GTCACCTTTTAAGATG-----TCAGAAATAATG 271
DB 484 SerLysLeuLysLeuValGlnPheValLeuAsnGluAlaLeuGlnValSerGluIleAla 503
QY 272 ATT-----GATGCT----- 274
DB 504 IleLysAspLeuProTyrSerLeuLeuAlaValPheSerTrpLeuGluTyrIleValArg 523
QY 275 -----GATGCT----- 280
DB 524 ValGlyTrpThrProAspAlaLysTyrValTrpValGlnGlyLeuAspArgLysGlnGln 543
QY 280 ----- 280
DB 544 ArgLeuAspValIleLeuIleProLeuAspAsnPheCysGluSerTyrSerSerGlnVal 563
QY 280 ----- 280
DB 564 SerThrProThrAspSerIleGlyAspHisSerTrpArgSerLeuTyrSerArgThrIle 583
QY 280 ----- 280
DB 584 ThrProLeuGlnValIleTyrThrGluArgSerAspSerTrpIleAsnValHisAspMet 603
QY 280 ----- 280
DB 604 LeuHisPheLeuAspLeuThrGluThrSerValThrPheLeuTrpAlaSerGluGluThr 623
QY 280 ----- 280
DB 624 GlyPheArgHisLeuTyrLeuValThrAlaSerLeuLeuSerGlnAlaAsnGlyGln 643

```
OY 281 -----GAAGGA 286
Db 644 ProAspProGlySerValGlyAlaGlnProSerPheValAspLeuSerAlaLeuGlnPro 663
OY 287 AGGATCAT----- 295
Db 664 ArgIleLeuAsnLysValAlaLeuThrSerGlyGluTyrGluValLeuAlaArgAsnLeu 683
OY 296 ---GTTGATCAAGTCAGAGGCTGTATATTTTGAAGCACCACAAAGACTCCCTTTTAGAG 352
Db 684 TrpValAspLysAlaAsnLysLeuValTyrPheValGlyLeuArgAspThrProLeuGlu 703
OY 353 CATCACTGTACGTACTGATCTACGTACGTAAATCCTGGAGAGGTGACAAAGCTGACTGACCGT 412
Db 704 LysHisLeuTyrValValSerLeuGluArgProGluHisIleArgLeuLeuThrGluPro 723
OY 413 GGTACTCATCTCTTTC-----TGCAATCAGT 439
Db 724 GlyTyrSerTyrLeuValGluPheAspHisPheAsnAspLeuPheProIlePheSer 743
OY 440 CAGCACTGTGACTCTCTTTATAGTAAGTATAGTAAC---CAGAAAGATCCACACTGTGTG 496
Db 744 GlnGlnCysLysLeuMetLeuValTyrCysAsnIleGlnArgLeuProSerCysLys 763
OY 497 TCCTTTTACAAGCTATCAAGTCTCAAGATGACCAACTTGCACAAACAAAGGAATTTTGG 556
Db 764 ValMet-----ArgValAsnGlnThrCysSerAsnGlyGlyValAsn 777
OY 557 GCCACCATTTTGGATTCAGCAGGT-----CCTCTCTCTGACTAT 595
Db 778 GlyIleGlnIleSerLeuValGlyTyrLeuHisGluGlyLysProGluProGlnTyr 797
OY 596 ACTCTCTCAGAAATTTCTCTTTTGAAGTACTACTGATTCATGATTCATGATGGATGCTC 655
Db 798 Cys---ProGlnIlePheSerProGlnLeuProSerGlyAspIleValTyrAlaMetVal 816
OY 656 TACAAGCTCATGATCTACACCTGGAAGAAATATCTACTCTGCTGCTTCATATATGTT 715
Db 817 PheLysProHisAsnPheGlnLeuGlyValLysTyrProThrValLeuAsnValTyrGly 836
OY 716 GTCTCTCAGGTGCGATGTTGTAATTCGGTTTAAAGGAGTCAAGTATTCCTCGTTGAAT 775
Db 837 GlyProGluValGlnThrValAsnAsnThrPheLysGlyLysHisGlnLeuArgMetHis 856
OY 776 ACCCTAGCTCTCTAGTTATGCTGTGTAGTGATGATGACACAAAGGGATCTCTGTCACCGA 835
Db 857 MetLeuAlaIleGlnGlyTyrCysValIleCysIleAspSerArgGlySerArgHisArg 876
OY 836 GGGCTTAATTTGAAGCGCTTTAAATATAAAATGGTCAATAGAAATTCACCATCAG 895
Db 877 GlyLysArgPheGluSerHisIleArgGlyArgMetGlyGlnValGluLeuThrAspGln 896
OY 896 GTGAAGACTCTCAATATCTAGCTTCTCGATATGATTTTCATTCAGTCTGATGCTGTGGC 955
Db 897 ValAspAlaLeuArgSerLeuSerAspGlnLeuGlyTyrIleAspMetAspArgValAla 916
OY 956 ATCCACGGCTGGTCTTATGGAGGATACCTCTCCCTGATGGCATTAATGACAGAGGTCAGAT 1015
Db 917 IleHisGlyTyrSerTyrGlyTyrLeuSerLeuMetGlyLeuValGlnTyrProLys 936
OY 1016 ATCTTCAGGGTGTCTATTGCTGGGCCCGGACGTCCTGCTGATCTCTTATGATACAGGA 1075
Db 937 IlePheLysValAlaIleAlaGlyAlaProValThrAsnTyrGluTyrTyrAspThrGly 956
OY 1076 TACACG 1081
Db 957 TyrThr 958
RESULT 9
ID Q9VC19 DROME PRELIMINARY; PRT; 1113 AA.
AC Q9VC19
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
```

```
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE CG3744-PB, isoform B.
GN Name=CG3744; ORFNames=CG3744;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
NCBI_TaxID=7227;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers V.-H.C., Blasej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Neilson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Frankoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Furler C.J., Evangelista C.C., Ferraz C., Ferrera S., Fleischmann W.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Houston K.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostalka M., Kaul S., Karpman G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Markulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinstock G., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E.C., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zhang L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Zhang L.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C., Smith H.O.,
RT "The genome sequence of Drosophila melanogaster.";
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426065; PubMed=12537568;
RA Celniker S.E., Wheeler D.A., Krommiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA George R.A., Hoskins R.A., Laverly T., Muzny D.M., Nelson C.R.,
RA Paclet J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT "Finishing a whole-genome shotgun: release 3 of the Drosophila
RL melanogaster euchromatic genome sequence.";
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426070; PubMed=12537573;
RA Kaminker J.S., Bergman C.M., Krommiller B., Carlson J.W., Svirskas R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celniker S.E.;
RT "The transposable elements of the Drosophila melanogaster euchromatin:
RL a genomics perspective.";
RL Genome Biol. 3:RESEARCH0084.1-RESEARCH0084.20(2002).
RN [4]
RP NUCLEOTIDE SEQUENCE.
```

RX MEDLINE=22426069; PubMed=12537572;
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.B.,
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
 RA Bettencourt B.R., Celisner S.E., de Grey A.D.N.J., Drysdale R.A.,
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
 RA Lewis S.E.;
 RA "Annotation of the Drosophila melanogaster euchromatic genome: a
 RT systematic review.";
 RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
 RN [5]
 RP NUCLEOTIDE SEQUENCE.
 RG Berkeley Drosophila Genome Project;
 RA Celisner S., Stapleton M., Pacleb J., Park S., Svirskas R., Smith E.,
 RA Hoskins R., Rubin G.;
 RT "Drosophila melanogaster release 4 sequence.";
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP NUCLEOTIDE SEQUENCE.
 RG FlyBase;
 RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AS003749; AAF56356.2; -; Genomic DNA.
 DR MEROPS; S09_016; -;
 DR Ensembl; CG3744; Drosophila melanogaster.
 DR FlyBase; FBgn0039240; CG3744.
 DR GO; GO:0003824; F: catalytic activity; IEA.
 DR GO; GO:0008236; F: serine-type peptidase activity; IEA.
 DR GO; GO:0006508; P: proteolysis and peptidolysis; IEA.
 DR InterPro; IPR001375; Peptidase_S9.
 DR InterPro; IPR002469; Peptidase_S9B.
 DR InterPro; IPR000379; Ser esters.
 DR Pfam; PF00930; DPPIV_N; 1.
 DR Pfam; PF00326; Peptidase_S9; 1.
 SQ SEQUENCE 1113 AA; 123681 MW; 20857E3B212DF2E4 CRC64;

Alignment Scores:
 Pred. No.: 7.28e-55 Length: 1113
 Score: 718.50 Matches: 176
 Percent Similarity: 41.1% Conservative: 55
 Percent Local Similarity: 31.1% Mismatches: 116
 Query Match: 37.1% Indels: 215
 DB: 2 Gaps: 10

us-10-825-632-8 (1-1083) x Q9VC19_DROME (1-1113)

QY	5	GRAGATGCCAGATCAGTCGAGTCGCTACCTTTGTTCTCCAAAGAAATTTGATAGATAT	64
DB	469	ASpAspAlaLeuSerAlaGlyValProSerTyrValMetGlnGluGluPheSerArgTyr	488
QY	65	TCCTGGTATTGGTGTGTCACAAAGCTCAACCACTCCACGTCGTGGTAAATCTTTAGA	124
DB	489	GlnGlyPheTrpTrpGlnProHis-----SerAsnAspGlyIleTyrArg	503
QY	125	ATTCTATATGAAGAAATGATGAACTGAGTGGGAATATTATGTTACATCCCTATG	184
DB	504	IleValTyrGluGluValAspGluSerGluValSerValTyrPheProSerSerThr	523
QY	185	TTGGAACAGAGGAGGACAGTTCATTCCGTTATCTCTAAACAGGTACAGCAATCCCTAAA	244
DB	524	AlaMetHisGlyValAspGluTyrArgPheProArgThrGlySerProAsnAlaLys	543
QY	245	GTCACTTTTAAAGATG-----TCAGAAATAATG	271
DB	544	SerLysLeuLysLeuValGlnPheValLeuAsnGluAlaLeuGlnValSerGluIleAla	563
QY	272	ATT-----	274
DB	564	IleLysAspLeuProTyrSerLeuLeuAlaValPheSerTrpLeuGluTyrIleValArg	583
QY	275	-----GATGCT-----	280

DB	584	ValGlyTrpThrProAspAlaLysTyrValTrpValGlnGlyLeuAspArgLysGlnGln	603
QY	280	-----	280
DB	604	ArgLeuAspValIleLeuIleProLeuAspAsnPheCysGluSerTyrSerSerGlnVal	623
QY	280	-----	280
DB	624	SerThrProThrAspSerIleGlyAspHisSerTrpArgSerLeuTyrSerArgThrIle	643
QY	280	-----	280
DB	644	ThrProLeuGlnValIleTyrThrGluArgSerAspSerTrpIleAsnValHisAspMet	663
QY	280	-----	280
DB	664	LeuHisPheLeuAspLeuThrGluThrSerValThrPheLeuTrpAlaSerGluGluThr	683
QY	280	-----	280
DB	684	GlyPheArgHisLeuTyrLeuValThrAlaSerLeuLeuSerGlnAlaAsnGlyGln	703
QY	281	-----GAAGGA	286
DB	704	ProAspProGlySerValGlyAlaGlnProSerPheValAspLeuSerAlaLeuGlnPro	723
QY	287	AGGATCATTA-----	295
DB	724	ArgIleLeuAsnLysValAlaLeuThrSerGlyGluTrpGluValLeuAlaArgAsnLeu	743
QY	296	---GTTGATGAGTCAGAAAGCGTGGTATATTTGAAGGCACCAAGACTCCCTTTAGAG	352
DB	744	TrpValAspLysAlaAsnLysLeuValTyrPheValGlyLeuArgAspThrProLeuGlu	763
QY	353	CATCAGCTGACGTAGTACGTACGTAACTCTCGAGAGGTGACAGAGCTGACGTACCGT	412
DB	764	LysHisLeuTyrValValSerLeuGluArgProGluHisIleArgLeuGluPro	783
QY	413	GGCTACTCAGATCTTTCG-----TGCTATCAGT	439
DB	784	GlyTyrSerTyrLeuValGluPheAspAspHisPheAsnAspLeuPheProIlePheSer	803
QY	440	CAGCAGCTGTGACCTCTTTTATAAGTAAGTATAGTAAC---CAGAAGAAATCCACGTGTG	496
DB	804	GlnGlnCysLysLeuMetLeuValTyrCysAsnIleGlnArgLeuProSerCysLys	823
QY	497	TCCCTTTAAGCTATCAAGTCTCTGAGATGATGATGATGATGATGATGATGATGATGAT	556
DB	824	ValMet-----ArgValAsnGlnThrCysSerAsnGlyGlyValAsn	837
QY	557	GCCACCATTTTGGATTCAGCAGGT-----CCTCTTCTGACTAT	595
DB	838	GlyIleGlnLysSerLeuValGlyTyrLeuHisGluGlyGlyLysProGluProGlnTyr	857
QY	596	ACTCCTCCAGAAATTTCTTTTGAAGTACTACTGATTTACATTTGATTTGATTTGAT	655
DB	858	Cys---ProGlnIlePheSerProGlnLeuProSerGlyAspIleValTyrAlaMetVal	876
QY	656	TACAGCCTCATGATCTACAGCCTGGAAGAAATATCTCTACTGCTGCTGCTCATATGCT	715
DB	877	PheLysProHisAsnPheGluLeuGlyValLysTyrProThrValLeuAsnValTyrGly	896
QY	716	GGTCTCAGCTGAGTGGTGAATAATCGTTTAAAGGAGTCAAGTATTTCCGCTTGAAT	775
DB	897	GlyProGluValGlnThrValAsnAsnThrPheLysGlyLysHisGlnLeuArgMetHis	916
QY	776	ACCCTAGCTCTCTAGTGTATGTGGTTGTAGTGATAGACACAGGGGATCTCTGTCCCGA	835
DB	917	MetLeuAlaAlaGlnGlyTyrCysValIleCysIleAspSerArgGlySerArgHisArg	936
QY	836	GGGCTTAAATTTGAAGCGCTTTTAAATATAAATCGGTCAATAGAAATTCACCATCAG	895
DB	937	GlyLysArgPheGluSerHisIleArgGlyArgMetGlyGlnValGluLeuThrAspGln	956


```

601 ProThrProProProThrProIleAspAlaLeuAsgLeuCysSerValGlyTyrLeuThr 620
602 -----CCAGAAATTTCTCTTTTGAAGTACT 628
621 GluGlyGlyProSerGluAsnThrGlnTyrAsnProSerIleHisSerProGlnIleSer 640
629 ACTGGATTTACATTGATGGGATGCTTACAGCCTCATGATCTACAGCCTGGAAAGAAA 688
641 SerGlyAspValLeuTyrAlaMetValPheLysProHisAsnPheMetLeuGlyValLys 660
689 TATCTACTGCTGCTTCATATATGTTGCTCAGTGCAGTGGTGGTGAATAATCGGTTT 748
661 TyrProThrValLeuAsnValTyrGlyGlyProGluValGlnThrValSerAsnThrPhe 680
749 AAGAGGACTCAAGTATTTCCGCTTGAATACCCCTAGCTCTCTAGGTTATGTTGTAGTG 808
681 LysGlyMetArgGlnLeuArgMetHisMetLeuAlaSerGlnGlyTyrCysValIleCys 700
809 ATAGACAAACAGGGGATCTCTGTCACCGAGGCTTAAATTTGAAGGCGCTTTAAATATAAA 868
701 ValAspSerArgGlySerArgHisArgGlyValGluPheGluSerTyrIleArgCysArg 720
869 ATGGGTCATAGAAATGACATGACATGAGTGGAGGACTCCAAATATCTAGCTTCTCGATAT 928
721 MetGlyThrValGluLeuSerAspGlnValGluValLeuArgIleLeuAlaAspGlnLeu 740
929 GATTCAATGACTAGTCTGCTGGCATCCAGCGTGGTCTCTATGGAGGATACCTCTCC 988
741 GlyTyrIleAspMetAspArgValAlaIleHisGlyTyrSerTyrGlyTyrLeuSer 760
989 CTGATGCGATTAATGACAGAGTGCAGATATCTTCAGGGTTGCTATTGCTGGGGCCCCAGTC 1048
761 LeuMetGlyLeuValGlnTyrProGluIlePheLysValSerIleAlaGlyAlaProVal 780
1049 ACTCTGGATCTCTTATGATACAGGATACAG 1081
781 ThrSerTrpGluTyrTyrAspThrGlyTyrThr 791

RESULT 11
Q5TTK8_ANOGA PRELIMINARY; PRT; 740 AA.
AC Q5TTK8;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE ENSANGP0000026132 (Fragment).
GN ORFNames=ENSANGG0000014037;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Culicidae;
OC Anophelinae; Anopheles.
OX NCBI_TaxID=180454;
[1]
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=PEST;
RC The Anopheles gambiae Sequence Committee;
RG The Anopheles gambiae re-annotation.;
RL Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
[2]
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=PEST;
RC The Anopheles gambiae Sequence Committee;
RG The Anopheles gambiae re-annotation.;
RL Submitted (APR-2004) to the EMBL/GenBank/DDBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAB01008879; EAL40717.1; -; Genomic DNA.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001375; Peptidase_S9.
DR InterPro; IPR002469; Peptidase_S9B.
```

```

DR InterPro; IPR000379; Ser_estr.
DR Pfam; PF00930; DPPIV N; 1.
DR Pfam; PF00326; Peptidase_S9; 1.
FT NON TER 1
SQ SEQUENCE 740 AA; 84151 MW; D59F131BBA095CCB CRC64;

Alignment Scores:
Pred. No.: 8,01e-52 Length: 740
Score: 684.00 Matches: 165
Percent Similarity: 42.1% Conservative: 67
Best Local Similarity: 29.9% Mismatches: 111
Query Match: 35.3% Indels: 208
DB: 2 Gaps: 9

US-10-825-632-8 (1-1083) x Q5TTK8_ANOGA (1-740)

QY 5 GAAGATCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAAGAGAAATTTGATGATAT 64
DQ 110 GluAspProLeuThrAlaGlyValProSerTyrValMetGlnGluPheSerArgTyr 129
QY 65 TCTGGCTATTGGTGGTGTCCAAAAGCTGAAACAACTCCCGAGTGGTGTAA 115
DQ 130 GlnGlyPheTrpTrpGlnProGlu 144
QY 116 -----ATTCTTGAATTTCTATATGAAGAAATGATGAATCTGAGTG 157
DQ 145 AspValIleLeuAspGluValTyrArgIleValTyrGluGluValAspGluSerAspVal 164
QY 158 GAAATTTATTTATGATATCCCTATGTTGGAAACAAAGGAGGAGGAGATTCATTCGTTAT 217
DQ 165 SerLeuTyrThrPheProSerSerGlnSerAlaGlyArgAspTyrGluGluTyrArgPhe 184
QY 218 CCTAAACAGGTACAGCAAAATCCTAAA 262
DQ 195 ProArgAlaGlyThrProAsnAlaLysSerLysLeuLysValGlnPheArgLeuSer 204
QY 263 GAA 265
DQ 205 GluAsnLeuArgIleThrAspValCysIleLysGluLeuGlnCysProLeuThrPheAla 224
QY 265 ----- 265
DQ 225 PheProTrpLeuGluTyrIleValArgValGlyTyrThrProAspSerArgTyrValTrp 244
QY 266 -----ATAATGATTGATCTGAA 283
DQ 245 AlaGlnLeuLeuAspArgProGlnGlnArgLeuGluLeuValLeuLeuProValAspAsn 264
QY 283 ----- 283
DQ 265 PheCysGluIleTyrSerSerSerSerSerArgProProAlaGlyAlaHisArgLysSerThr 284
QY 283 ----- 283
DQ 285 ThrArgProLeuGlnValIleTyrThrGluThrSerSerSerTrpValAsnValHisAsp 304
QY 283 ----- 283
DQ 305 ValLeuGlnPheValGluLeuSerGluGlnGluValThrPheLeuTrpAlaSerGluGlu 324
QY 283 ----- 283
DQ 325 SerGlyPheArgHisLeuTyrLeuValThrSerLysValSerSerGlyGlyValGlyAla 344
QY 284 -----GGAAGGATCAT 295
DQ 345 ThrAspHisSerLeuProSerMetAlaCysIleGlySerThrLeuValAlaArgIleVal 364
QY 296 -----GTTGAT 301
DQ 365 GlnLysValThrLeuThrGlyGlyAspTyrGluValLeuGlyArgAsnValTrpTyrAsp 384
QY 302 GAAGTCAGAGGCTGGTATATTTTGAAGGACCAAGACTCCCTTTTAGAGCATCACCTG 361
```

```
Db 385 ArgAlaArgGlnLeuValTyrPheMetGlyLeuAArgGluThrProLeuGluGlyHisLeu 404
QY 362 TAGCTAGTCAAGTTACGTAATCTCGAGAGGTGACAGCTGACTGACCGTGGCTACTCA 421
Db 405 TyrValValSerLeuAlaGlnProAenGlnLeuAArgLeuLeuThrMetProGlyTyrSer 424
QY 422 CATCTCTGTCGATCAGTCAGCTGCTCTCTTTATAGTATAGTATAGTAAC----- 475
Db 425 -----PheThrValGluPheAenAseValSer 433
QY 476 -----CAGAAAGATCCACATGCTGTGTGCTCTTACAAGCTATCAAGTCTCGAAGATGAC 529
Db 434 ValPheGluArgGlyGlyAlaCysProAenThrAlaAenGlnAenGlyCysSerHisGly 453
QY 530 CCACTTGGCAAAACAAGAAATTTGGCCACCAATTTTGGATTGAGTTCAGCAGGTCTCTTCCT 599
Db 454 ProThrProProProThrProLeuAspAlaLeuAArgLeuCysSerValGlyTyrLeuThr 473
QY 590 GACTATATCTCTCT-----CCAGAAATTTTCTCTCTTTTGAAGTACT 628
Db 474 GluGlyGlyProSerGluAenThrGlnTyrAenProSerIleHisSerProGlnIleSer 493
QY 629 ACTGGATTTACATTTGATGGATGCTCTACAGGCTCATGATCTACAGCTTGAAGAA 688
Db 494 SerGlyAspValLeuTyrAlaMetValPheLysProHisAenPheMetLeuGlyValLys 513
QY 689 TATCTTACTGTGCTGTTCATATATGCTGCTCAGGTGCAGTTGGTGAATATCGTTT 748
Db 514 TyrProThrValLeuAenValTyrGlyGlyProGluValGlnThrValSerAenThrPhe 533
QY 749 AAGAGGATCAAGTATTTCCGTTGAATACCTACCTGCTCTAGGTTATGTTGTTAGTG 808
Db 534 LysGlyMetArgGlnLeuArgMetHisMetLeuAlaSerGlnGlyTyrCysValIleCys 553
QY 809 ATAGACAACAGGGATCTCTGTCACCGAGGCTTAAATTTGAGCGCGCTTAAATATAA 868
Db 554 ValAspSerArgGlySerArgHisArgGlyValGluGluGluSerTyrIleArgCysArg 573
QY 869 ATGGTCAATAGAAATTCAGATCAGTGTGGAAGGATCTCCAATATCTAGTCTTCGATAT 928
Db 574 MetGlyThrValGluLeuSerAseGlnValGluValLeuAArgIleLeuAlaAseGlnLeu 593
QY 929 GATTCATTGATTCATGATCGTGTGGGCATCCACGGCTGCTCTAGGAGGATACCTCTCC 988
Db 594 GlyTyrIleAsePheMetAspArgValAlaIleHisGlyTyrSerTyrGlyGlyTyrLeuSer 613
QY 989 GTGATGGCAATTAATCAGAGGTGATATCTTACGGGTTGCTATTGCTGGGGCCCGATC 1048
Db 614 LeuMetGlyLeuValGlnTyrProGluIlePheLysValSerIleAlaGlyAlaProVal 633
QY 1049 ACTCTGTGATCTCTTATGATACAGGATACACG 1081
Db 634 ThrSerTrpGluTyrTyrAspThrGlyTyrThr 644

RESULT 12
Q7PTT8 ANOGA
ID Q7PTT8 ANOGA PRELIMINARY; PRT; 621 AA.
AC
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE ENSANGP0000015447 (Fragment).
GN ORFNames=ENSANGG0000012958;
OS
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae;
OC Anophelinae; Anopheles.
OX NCBI_TaxID=180434;
RN
RP NUCLEOTIDE SEQUENCE.
RC The Anopheles gambiae Sequence Committee;
Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
preliminary data.
DR EMBL; AA0801008287; EAA03335.3; -; Genomic_DNA.
DR MEROPS; S09.016; -.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001375; Peptidase_S9.
DR InterPro; IPR002469; Peptidase_S9B.
DR InterPro; IPR000379; Ser_esters.
DR Pfam; PF00930; DPVIV_N; I.
DR Pfam; PF00326; Peptidase_S9; 1.
FT NON_TER 1 621
FT NON_TER 621 621
SQ SEQUENCE 621 AA; 70546 MW; F6EA8463A343BBF3 CRC64;

Alignment Scores:
Pred. No.: 4,99e-47 Length: 621
Score: 630.00 Matches: 150
Percent Similarity: 49.6% Conservative: 54
Best Local Similarity: 36.5% Mismatches: 117
Query Match: 32.5% Indels: 90
DB: Gaps: 10

US-10-825-632-8 (1-1083) x Q7PTT8 ANOGA (1-621)
QY 65 TCTGGCTATTGGTGGTGTCCTCA-----AAAGCTGAACAACACTCCAGTGGTGTAAATTT 118
Db 134 SerSerGlyTyrArgSerProLeuAspLysSerThrThrArgPro----- 148
QY 119 CTTAGAATCTATATGAGAAATATGATGAATCTGAGGTGGAAATTTATTCATGTTACATCC 178
Db 149 LeuGlnValIleTyrThrGluThrSerSerSerTyrValAenValHisAseValLeuGln 168
QY 179 -----CCTATGTTGGAAACAAGAGGGCAGATTCATTC 211
Db 169 PheValGluLeuSerGluGlnGluValThrPheLeuTyrPalaSerGluGluSerGlyPhe 188
QY 212 CGT-----TATCCTAAAACAGGTACAGCAATCTCT----- 241
Db 189 ArgHisLeuTyrLeuValThrSerSerLeuSerProAenGluValSerSerGlyGlyAla 208
QY 241 ----- 241
Db 209 GlyAlaThrAspHisSerLeuProSerMetAlaCysIleGlySerThrLeuValAlaArg 228
QY 242 -----AAAGTCACCTTTAAGATGTCAGAAATATGATGATGCTGAAGNAGG--- 289
Db 229 IleValGlnLysValThrLeuThrGlyClyAsp-----TrpGluValLeuGlyArgAen 246
QY 290 ATCATAGTTGATGAAGTCAAGAGGTGGTATATTTGAAGCCACCAAGACTCCCTTTTA 349
Db 247 ValTrpTyrAspArgValArgGlnLeuValTyrPheMetGlyLeuAArgGluThrProLeu 266
QY 350 GAGCATCCTCTAGTGTAGTACCTAAATCCTGGAGAGGTGACCAAGGCTGACTGAC 409
Db 267 GluLysHisLeuTyrValValSerLeuAlaGlnProAenGlnLeuAArgLeuThrMet 286
QY 410 CGTGGCTACTCATATTCTTGCTGCATCAGTCAGCAGCTGCTGACTTCTTTATAAGTAAGTAT 469
Db 287 ProGlyTyrSerPheThrValGluPheAenAsePheCysThrLeuPheLeuGlnThrTyr 306
QY 470 AGTAACCAAGAAATCCACACTGTGTGCTCCCTTTACAAAGCTA-----TCAAGTCT 520
Db 307 CysAenIleSerThrLeuProSerTrpGluLeuValArgIleAlaHisAsePheSerAenThr 326
```

```
RT "Anopheles gambiae re-annotation.";
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
RN (2)
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
preliminary data.
DR EMBL; AA0801008287; EAA03335.3; -; Genomic_DNA.
DR MEROPS; S09.016; -.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001375; Peptidase_S9.
DR InterPro; IPR002469; Peptidase_S9B.
DR InterPro; IPR000379; Ser_esters.
DR Pfam; PF00930; DPVIV_N; I.
DR Pfam; PF00326; Peptidase_S9; 1.
FT NON_TER 1 621
FT NON_TER 621 621
SQ SEQUENCE 621 AA; 70546 MW; F6EA8463A343BBF3 CRC64;
```


QY 521 GAAGATGACCCAACTTGCACAAACAAAGAAATTTTGGGCCACCATTGTTGATTTCACGAGT 580
Db : : : : :
QY 327 AlaAenGlyAenGlyCys-----SerHisGly 335
QY 581 CTTCTTCTGACTATATCTCT----- 601
Db : : : : :
QY 336 ProThrProProProThrProLeuAspAlaLeuArgLeuCysSerValGlyTyrLeuThr 355
QY 602 -----CCAGAAATTTCTCTTTTGAAGAGTACT 628
Db : : : : :
QY 356 GluGlyGlyProSerGluAenThrGlnTyrAsnProSerIleHisSerProGlnIleSer 375
QY 629 ACTGATTTACATTTGATGGATGCTCTACAAGCCTCATGATCTACAGCTCGAAGAAA 688
Db : : : : :
QY 376 SerGlyAaspValLeuTyrAlaMetValPheLysProHisAenPheMetLeuGlyVallys 395
QY 689 TATCTCTCTGCTGCTTTCATATATGGTGGTCTCAGGTGAGTGGTGAATATCGGTTT 748
Db : : : : :
QY 396 TyrProThrValLeuAenValTyrGlyGlyProGluValGlnThrValSerAenThrPhe 415
QY 749 AAAGGAGTCAAGATTTTCCCGTTGAATACCTAGCTCTCTAGTTATGTGTTGTAGTG 808
Db : : : : :
QY 416 LysGlyMetArgGlnLeuArgMetHisMetLeuAlaSerGlnGlyTyrCysValIleCys 435
QY 809 ATAGACACAGGGGATCTGTGTCACGAGGGCTTAATTTCAAGCGCCCTTTAAATATAA 868
Db : : : : :
QY 436 ValAaspSerArgGlySerArgHisArgGlyValGluPheGluSerTyrIleArgArg 455
QY 869 ATGGGTCAATAGAAATGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATAT 928
Db : : : : :
QY 456 MetGlyThrValGluLeuSerAspGlnValGluValLeuArgIleLeuAlaaspGlnLeu 475
QY 929 GATTTTCATTGACTAGCTGTGGGCATCCACGGCTGGTCTTATGAGGATACCTCTCC 988
Db : : : : :
QY 476 GlyTyrIleAspMetAspArgValAlaIleHisGlyTyrSerTyrGlyGlyTyrLeuSer 495
QY 989 CTGATGGCATTAATGACAGGTGCAGATATCTCAGGTTGCTATTGCTGGGCCCCAGTC 1048
Db : : : : :
QY 496 LeuMetGlyLeuValGlnTyrProGluIlePheLysValSerIleAlaGlyAlaProVal 515
QY 1049 ACTCTGTGATCTCTCTATGATACAGGATACAG 1081
Db : : : : :
QY 516 ThrSerTrpGluTyrTyrAspThrGlyTyrThr 526

RESULT 13
Q54U01.DICDI
ID Q54U01.DICDI PRELIMINARY; PRT; 803 AA.
AC Q54U01; 2005 (T-EMBLrel. 31, Created)
DT 13-SEP-2005 (T-EMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (T-EMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=DD80205566;
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=AX4;
RA Eichinger L., Pachebat J.A., Gloeckner G., Rajandream M.A.,
RA Sugang R., Berriman M., Song J., Olsen R., Szafranski K., Xu Q.,
RA Tunggal B., Kummerfeld S., Madera M., Konfortov B.A., Rivero P.,
RA Bankier A.T., Lehmann R., Hamlin N., Davies R., Gaudet P., Fey P.,
RA Pilcher K., Chen K., Saunders D., Sodergren E., Davis P.,
RA Kerhornou A., Nie X., Hall N., Anjard C., Hemphill L., Bason N.,
RA Farbrother P., Desany B., Just E., Morio T., Rost R., Churcher C.,
RA Cooper J., Haydock S., van Driessche N., Cronin A., Goodhead I.,
RA Muzny D., Mourier T., Pain A., Lu M., Harper D., Lindsay R.,
RA Hauser H., James K., Quiles M., Mohan M.B., Saito T., Buchrieser C.,
RA Wardroper A., Felder M., Thangavelu M., Johnson D., Knights A.,
RA Loulseghe H., Mungall K., Oliver K., Price C., Quail M.A.,
RA Urushihara H., Hernandez J., Rabinowitsch E., Steffen D., Sanders M.,

RA Ma J., Kohara Y., Sharp S., Simmonds M., Spiegler S., Tivey A.,
RA Sugano S., White B., Walker D., Woodward J., Winckler T., Tanaka Y.,
RA Shaulsky G., Schleicher M., Weinstock G., Rosenthal A., Cox E.C.,
RA Chisholm R.L., Gibbs R., Loomis W.F., Platzner M., Kay R.R.,
RA Williams J., Dear P.H., Noegel A.A., Barrell B., Kuspa A.,
RT "The genome of the social amoeba Dictyostelium discoideum.";
RL Nature 0:0-0(2005).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AA01000074; EMBL66689.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 803 AA; 93187 MW; 33846B370C039FAE CRC64;

Alignment Scores:
Pred. No.: 7,88e-45 Length: 803
Score: 605.50 Matches: 156
Percent Similarity: 39.8% Conservative: 59
Best Local Similarity: 28.9% Mismatches: 122
Query Match: 31.2% Indels: 203
DB: Gaps: 13

US-10-825-632-8 (1-1083) x Q54U01.DICDI (1-803)
QY 14 AGATCAGCTGGAGTCGCTACCTTTGTTCTCCAAAGAAATTTGATAGATATTCGCGTAT 73
Db 189 AtgTyrAlaGlyAspIleGlyPheIleTyrAlaGluGluPheSerArgTyrThrGlyTyr 208
QY 74 TGTGTGTGTCCTCAAAAGCTGAACTCCAGTCGTGTGTA---ATTCTTAGAATTTCTA 130
Db 209 TrpTrpSerProIleValGlyThrCysVallysThrGlyLysProMetTyrThrIleCys 228
QY 131 TATGAAGAAATGATGAATCTCAGGTGGAAATTTATTCATGTATCATCCCTATGTTGGAA 190
Db 229 TyrLeuGluGluaspGluThrAsnValMetAspTyrHisIle-----ProThrSerAsp 246
QY 191 ACAAG---AGGCGAGATTCATTCGTTATCTCTAAACAGGTACAGCAAT----- 238
Db 247 LeuArgGlyLysThrThrGlnTyrLysTyrProLeuAlaGlyGluLysAsnSerIleCys 266
QY 239 -----CCTAAGTCACCTTTTAAAGATCTCAGAAATA 268
Db 267 LysValCysLeuValSerPheValLeuProThrArgThrPheGlnaspSerLysIle 286
QY 269 ATGATT----- 274
Db 287 GluIleValLysSerGluLeuPheaspLeuLysThrGlnPheProTrpAlaGluTyrIle 306
QY 274 ----- 274
Db 307 ThrArgAlaGlyTrpThrProaspGlyHisSerIleTyrLeuGlnLeuLeuaspArgLys 326
QY 274 ----- 274
Db 327 GlnGlnHisLeuAlaLeuValMetValProLeuHisValPheAlaGluaspTyr-SerSer 346
QY 274 ----- 274
Db 347 SerSerSerSerValLysSerIleProLysLeuProValLeuIleGluGluThrThr 366
QY 275 -----GATGCTGAAGAGGATC 292
Db 367 SerValTrpIleAsnIleGluPheSerPheGlnPheLeuLysSerIleGluAsnGlnLeu 386
QY 293 ATA----- 295
Db 397 IleTrpSerAsnGluGlnSerGlyTyrArgHisLeuTyrLeuIleLysTrpAspLysAsn 406
QY 295 ----- 295
Db 407 PheThrAsnIleGlnSerThrThrProIleThrLeuSerThrCysAsnAspAsnAsp 426
QY 296 -----GTTGATCAAGTCAGAGGCTGGTA 319

```
Db 427 AsnAsnAenTrpMetValSerSerAspAspIleHisIleAspGluLysArgLysLeuVal 446
Qy 320 TATTTTGAAGGACCAAGACCTCCCTTTAGACATCCCTAGCTAGCTAGCTAGCTAGCTA 379
Db 447 TyrPheThrGlyThrLysAspThrCysLeuGluGlnHisLeuValThrArgPheAsp 466
Qy 380 AATCCT---GGAGAGTGCACAGGCTGACTGACCTGGCTGCTACTCAT---TCWTGCTGC 433
Db 467 LysProAsnSerGluIleLysArgLeuSerHisAlaAsnPheSerHisArgSerIleSer 486
Qy 434 ATCAGTCAGCAGCTGCTCCTTTTATAGTAACTAGTAACTAGTAACTAGTAACTAGTAA 484
Db 487 IleSerSerAsnPhelyLysPheIleThrThrThrThrThrThrThrThrThrThrThr 506
Qy 485 -----CCA 487
Db 507 ThrGluValPheAspLeuIleTyrAsnAspAspAsnAsnAspAsnAspAspIleTyrPro 526
Qy 488 CACTGTGTGCTCCTTTTACAGCTATCAAGTCTCTGAAGATGACCCCACTTGCACAAACAA 547
Db 527 IleValLysSerSerPhePheIleAsnAspAspAspAspAspAspAspAspAspLysLys 546
Qy 548 GAATTTTGGGCCACCATTTTGGATTTCAGCAGGCTCTCTCTCTGCTACTATCTCTCTCAGAA 607
Db 547 LysIleAsnIleAsnIle-----ProLys 554
Qy 608 ATTTCTCTTTTGAAGTACTACTGATTTTATGATGATGATGATGATGATGATGATGATGAT 667
Db 555 IlePheAsnPhelyLysAsnSerLysGlyValThrIleTyrGlyGlnTyrThrLeuProSer 574
Qy 668 GATCTACAGCCTGGAAGAAATATCTACTGCTGCTTATATATATATATATATATATATATAT 727
Db 575 AspTyrSerLysAspLysLysTyrProThrValValTyrValTyrGlyGlyProHisVal 594
Qy 728 CAGTTGGTGAATAATTCGGTTTAAAGGATCAAGTATTTCCGCTTGAATACCTAGCTCTCT 787
Db 595 GlnIleValArgAsnGlnTyrAsnTyrIleLysGlnHisTyrThrAsn-----610
Qy 788 CTAGTTTATGTTGTAGTATAGACAGACAGGAGTCTCTGTCACCGAGGCTTAAATTT 847
Db 611 PheGlyPheIleGlnValMetIleAspAsnValGlySerAlaAsnArgGlyLeuGluPhe 630
Qy 848 GAAGGCGCTTTAAATATATAATGCTCAATGAATTCACCATCAGTGTGAAGGATC 907
Db 631 GluSerHisIleArgGluLysMetGlyGlnValGluIleGlyAspGlnValGluGlyIle 650
Qy 908 CAATATCTAGCTCTCGATATGATTC-----ATTGACTTATGATGCTGTGGCCATCCAC 961
Db 651 AsnTyrLeuValGlyAsn---AspIleValSerIleAspValAsnArgIleAlaIleSer 669
Qy 962 GCCTGCTCTATGAGATACCTCTCCTGATGCATTAAATGACAGAGTTCAGATATCTTC 1021
Db 670 GlyTrpSerTyrGlyTyrAsnSerLeuMetAlaIleSerGlnArgProAspValPhe 689
Qy 1022 AGGTTGCTATGCTGGGCGCCAGTCACCTCTGATGATCTCTATGATACAGGATACACG 1081
Db 690 LysIleAlaValCysGlyAlaProValSerAspTrpArgLeuTyrAsnThrGlyTyrThr 709

RESULT 14
Q5TXJ2_ANOGA PRELIMINARY; PRT; 557 AA.
AC Q5TXJ2;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DE ENSANGP0000029249 (Fragment)
GN ORFNames=ENSANGG0000012958;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Culicidae;
OC Anophelinae; Anopheles.
OX NCBI_TaxID=180454;
```

```
RN NUCLEOTIDE SEQUENCE.
RC STRAIN-PEST;
RG The Anopheles gambiae Sequence Committee;
RT "Anopheles gambiae re-annotation.";
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
[2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN-PEST;
RG The Anopheles gambiae Sequence Committee;
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAB01008287; EAL42106.1; -!- Genomic_DNA.
DR GO; GO:0003824; Fcatalytic activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001375; Peptidase_S9.
DR InterPro; IPR002469; Peptidase_S9.
DR InterPro; IPR000379; SerProts.
DR Pfam; PF00910; DPPIV_N; I.
DR Pfam; PF00326; Peptidase_S9; 1.
DR NON_TER 557
FT NON_TER 557
SQ SEQUENCE 557 AA; 62685 MW; 08AAA0944ACD0808 CRC64;

Alignment Scores:
Pred. No.: 1.69e-44 Length: 557
Score: 601.50 Matches: 129
Percent Similarity: 63.6% Conservative: 42
Best Local Similarity: 48.0% Mismatches: 79
Query Match: 31.0% Indels: 19
DB: 2 Gaps: 5

US-10-825-632-8 (1-1083) x Q5TXJ2_ANOGA (1-557)
Qy 299 GATGAAGTCAGAGGCTGGTATATTTTGAAGGACCAAGACTCCCTTTAGAGACTCAC 358
Db 231 Aspa-gvalArgGlnLeuValTyr-PheMetGlyLeuArgGluThrProLeuGluLysHis 250
Qy 359 CTGTAGCTAGTCAGTACGTAAATCCCTGGAGAGTGACAAAGGCTGACTGACCTGGCTAC 418
Db 251 LeutyValValSerLeuAlaGlnProAsnGlnLeuArgLeuLeuThrMetProGlyTyr 270
Qy 419 TCACATTTCTTGTGTCATCAGTCAGCAGCTGTGACTTCTTT---ATAAGTAAGTATAGTAC 475
Db 271 SerPheThrVal-----GluPheAsnAspLeuValArgIleAlaHisAspSerAsn 287
Qy 476 CAGAGAATCCACACTGTGTGCTCTTACAGCTATCAAGCTCTGAAAGATGACCAACT 535
Db 288 ThrAlaAsnGlyAsnGlyCysSer-----HisGlyProThrProProProThr 303
Qy 536 TGCACAAACAAAGGAATTTTGGGCCACCATTTTGGATTTCAGCAGCTCTCTTCTGACTAT 595
Db 304 -----ProIleAspAlaLeuArgLeuCysSerValGlyTyrLeuThrGluGly 319
Qy 596 ACTCCT-----CCAGAAATTTCTCTTTTGAAGTACTACTGGA 634
Db 320 GlyProSerGluAsnThrGlnTyrAsnProSerIleHisSerProGlnIleSerSerGly 339
Qy 635 TTTCAATTTGATGGATGCTCTACAGCCTCATGATCTACAGCCTGGAAGAAATATCTCT 694
Db 340 AspValLeuTyrAlaMetValPheLysProHisAsnPheMetLeuGlyValLysTyrPro 359
Qy 695 ACTGCTGCTCATATATATGTTGCTCTCAGTGCAGTGTGTAATATCGTTTAAAGGA 754
Db 360 ThrValLeuAsnValTyrGlyProGluValGlnThrValSerAsnThrPheLysGly 379
Qy 755 GTCAGATTTTCCGCTTGAATACCTTCTCTAGTGTATGTTGTTGTTGTTGTTGTTGTTGTTG 814
Db 380 MetArgGlnLeuArgMetHisMetLeuAlaSerGlnGlyTyr-CysValIleCysValAsp 399
```

5 GAAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAAGAAGAAATTGATAGATAT 64

```

Db      524 ProHisAlaGlnargValMetLysAanTrpHisSerProSerGlu-----ArgThr 540
Qy      779 CTAGCCCTCTCTAGTTTATGTTGTAGTAGACACAGGGATCCTGTACCCGAGG 838
Db      541 TyrLeuGluAlaGlyTyrValIlePheLysLeuAspAenargGlySerGlyAsnArgSer 560
Qy      839 CTTAAATTTGAAGCGCCCTTAAATATATAAATGGTCAAATAGAAATTGACGATCAGGTG 898
Db      561 AlalysPheMetArgAlaLeuAspArgLysLeuGlyThrValGluValGluAspGlnLeu 580
Qy      899 GAAGGACTCCAATATCTAGCTTCTCGATATGATTTCATTGACTTAGATCGTGTGGGCATC 958
Db      581 LeuGlyAlalysPheLeuAlaSerGln---ProTyrValAspAlaAspLysLeuGlyVal 599
Qy      959 CACGGCTGTCCTATGGAGGATACCTCTCCCTGATGGCAATTAATGCAGAGTCAGATATC 1018
Db      600 MetGlyTrpSerTyrGlyGlyPheMetAlaLeuMetLeuLeuThrAlaGluAsnThrPro 619
Qy      1019 TTCAGGTTGCTATTCTCGGCCCCAGTCACCTCTGTGGATCTTCTATGATACAGGATAC 1078
Db      620 PheLysAlaGlyAlaAlaGlyAlaProProThrGluTrpSerLeuTyrAspThrAlaTyr 639
Qy      1079 ACG 1081
Db      640 Thr 640

```

Search completed: May 2, 2006, 01:34:05
 Job time : 274.427 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: May 2, 2006, 01:34:19 ; Search time 6.63677 Seconds
(without alignments)
2698.234 Million cell updates/sec

Title: US-10-825-632-8

Perfect score: 1938

Sequence: 1 ggaagaagatgccagatcag.....tatgatacagatacacgga 1083

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 1144120

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp
-Q=/abs/ABSSWEB.spool/US10825632/runat_01052006_105950_3311/app.query.fasta.1
-DB=Issued_Patents_AA -QFIX=fastan -SUFFIX=n2p.ra1 -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-HOST=abs802p -USER=US10825632@cgn_1_131@runat_01052006_105950_3311
-NCPU=6 -ICPU=3 -NO MMAP -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued_Patents_AA:*
1: /cgn2_5/ptodata/1/iaa/5 COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/6 COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/H COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/pCTUS COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/RE COMB.pep:*
6: /cgn2_5/ptodata/1/iaa/backfiles.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1933	99.7	360	2	US-10-070-464-7
2	1836.5	94.8	882	2	US-09-976-674-1
3	1836.5	94.8	882	2	US-10-070-464-1
4	1226	63.3	690	2	US-09-976-674-7
5	1220.5	63.0	863	2	US-09-976-674-3
6	1220.5	63.0	892	2	US-09-976-674-23
7	1220.5	63.0	892	2	US-09-976-674-27
8	1209	62.4	661	2	US-09-976-674-11
9	1201	62.0	879	2	US-09-976-674-33
10	1201	62.0	879	2	US-09-976-674-35
11	1199.5	61.9	658	2	US-09-976-674-19
12	1132.5	58.4	832	2	US-09-976-674-29

13	1132.5	58.4	832	2	US-09-976-674-31	Sequence 31, Appl
14	1113	57.4	819	2	US-09-976-674-37	Sequence 37, Appl
15	1113	57.4	819	2	US-09-976-674-39	Sequence 39, Appl
16	1086.5	56.1	310	2	US-09-794-236-4	Sequence 4, Appl
17	1086.5	56.1	310	2	US-10-070-464-3	Sequence 3, Appl
18	931.5	48.1	613	2	US-09-976-674-21	Sequence 21, Appl
19	860.5	44.4	465	2	US-10-070-464-5	Sequence 5, Appl
20	584	30.1	358	2	US-09-976-674-13	Sequence 13, Appl
21	444	22.9	981	2	US-09-902-540-16812	Sequence 16812, A
22	403.5	20.8	159	2	US-09-270-767-32313	Sequence 32313, A
23	403.5	20.8	732	2	US-09-518-550-30	Sequence 30, Appl
24	369	19.0	710	2	US-09-518-550-28	Sequence 28, Appl
25	352	18.2	518	2	US-09-976-674-25	Sequence 25, Appl
26	347.5	17.9	771	2	US-09-462-284-2	Sequence 2, Appl
27	347.5	17.9	771	2	US-09-079-592-2	Sequence 2, Appl
28	344	17.8	676	2	US-09-518-550-42	Sequence 42, Appl
29	344	17.8	723	2	US-09-518-550-29	Sequence 29, Appl
30	329	17.0	737	2	US-09-902-540-11421	Sequence 11421, A
31	315	16.3	755	4	PCT-US93-07923-3	Sequence 3, Appl
32	315	16.3	759	4	PCT-US93-07923-2	Sequence 2, Appl
33	315	16.3	766	1	US-08-230-491A-3	Sequence 3, Appl
34	315	16.3	766	1	US-08-619-280A-3	Sequence 3, Appl
35	315	16.3	766	1	US-08-940-391-3	Sequence 3, Appl
36	315	16.3	766	2	US-09-794-236-1	Sequence 1, Appl
37	315	16.3	766	2	US-10-002-593-6	Sequence 6, Appl
38	315	16.3	766	2	US-09-949-016-6146	Sequence 6, Appl
39	315	16.3	766	2	US-09-265-606-3	Sequence 3, Appl
40	315	16.3	766	2	US-10-423-714-6	Sequence 6, Appl
41	315	16.3	775	2	US-09-949-016-10450	Sequence 10450, A
42	309	15.9	766	2	US-09-518-550-27	Sequence 27, Appl
43	287.5	14.8	574	2	US-09-902-540-11805	Sequence 11805, A
44	280.5	14.5	818	2	US-09-462-845-3	Sequence 3, Appl
45	280.5	14.5	818	2	US-10-402-312-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1

US-10-070-464-7
; Sequence 7, Application US/10070464
; Patent No. 6881564
; GENERAL INFORMATION:
; APPLICANT: ABBOTT, Catherine Anne
; APPLICANT: GORRELL, Mark Douglas
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES
; FILE REFERENCE: GH-007
; CURRENT APPLICATION NUMBER: US/10/070,464
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: PCT/AU00/01085
; PRIOR FILING DATE: 2000-09-11
; PRIOR APPLICATION NUMBER: AU PQ5709
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: AU PQ2762
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 360
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-070-464-7

Alignment Scores:
Pred. No.: 3.13e-220 Length: 360
Score: 1933.00 Matches: 360
Percent Similarity: 100.0% Conservative: 0
Best local Similarity: 100.0% Mismatches: 0
Query Match: 99.7% Indels: 0
DB: 2 Gaps: 0

US-10-825-632-8 (1-1083) x US-10-070-464-7 (1-360)

QY 2 GAAGAAGATGCCAGATCAGCTGGAGTCGCTTGTCTTCCAAAGAAATTGTGATAGA 61

```
Db 1 GluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGluPheAspArg 20
Qy 62 TATTCTGGCTATTGGTGTCCAAAGCTGAACCACTCCAGTGGTGTGAATTCCT 121
Db 21 TyrSerGlyTyrTrpCysPheValGluThrThrProSerGlyGlyLysLeu 40
Qy 122 AGAATCTATATGAAGAAATGATGATCTGAGTGGAAATATATCATGTTACATCCCT 181
Db 41 ArgLeuLeuTyrGluGluAsnAspGluSerGluValGluLeuIleHisValThrSerPro 60
Qy 182 ATGCTGGAACAGGAGGACATTCATCCCTGATTCCTAAACAGGTACAGCAATCCT 241
Db 61 MetLeuGlnThrArgArgAlaAspSerPheArgTyrProLysThrGlyThrAlaAsnPro 80
Qy 242 AAGTCACTTTAAGATGTCAGAAATAATGATGATGCTGAAGGAAGGATCATAGTTGAT 301
Db 81 LysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIleIleValAsp 100
Qy 302 GAAGTCAGAGGCTGGTATATTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTG 361
Db 101 GluValArgArgLeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeu 120
Qy 362 TAGCTAGTCAGTTACGTAATCCTGGAGAGGTGCAAGGCTGACTGACCGTGGCTACTCA 421
Db 121 TyrValValSerTyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSer 140
Qy 422 CATCTCTGCTGATCAGTCAGTGTGACTTCTTTATAGTAAGTATAGTACACAGAAG 481
Db 141 HisSerCysCysIleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLys 160
Qy 482 AATCCACACTGTGTGCTCTTTACAAAGCTATCAAGTCTCTGAAGTACCCACTGTCGAA 541
Db 161 AsnProHisCysValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLys 180
Qy 542 ACAAGGAATTTTGGCCACCATTTTGGATTACAGCAGGTCTCTCTGCTGCTACTACTCT 601
Db 181 ThrLysGluPheTrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrPro 200
Qy 602 CAGAAATTTCTCTTTGAAGTACTGATGATTTACATGATGGGATGCTCTACAG 661
Db 201 ProGluIlePheSerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLys 220
Qy 662 CCTCATGATCTACAGCTGGAGAAATATCCTACTGTGCTGTTCATATATGCTGCTCT 721
Db 221 ProHisAspLeuGlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyPro 240
Qy 722 CAGTGCAGTTGGTGAATATCGTTTAAAGGAGTCAAGTATTCGCTTGAATACCTCA 781
Db 241 GlnValGlnLeuValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeu 260
Qy 782 GCCTCTAGGTTATGTTGTTAGTATAGTACACACAGGGGATCTCTGTCACCGAGGCTT 841
Db 261 AlaSerLeuGlyTyrValValValIleAspAsnArgGlySerCysHisArgGlyLeu 280
Qy 842 AAATTTGAAGCGCTTTAAATATAAATGGGTCAAAATAGAAATGACGATCAGGTGAA 901
Db 281 LysPheGluGlyAlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGlu 300
Qy 902 GGAATCAATATCTAGCTTCTCGATATGATTTCAATGATGCTGATGCTGGGATCCAC 961
Db 301 GlyLeuGlnTyrLeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyLeHis 320
Qy 962 GCCTGCTCCTATGAGGATACCTCTCCCTGATGCGATTAATCAGAGGTTCAGATATCTC 1021
Db 321 GlyTrpSerTyrGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePhe 340
Qy 1022 AGGTTGCTATTGCTGGGGCCAGTCACTCTGCTGGATCTCTATGATACAGGTACACG 1081
Db 341 ArgValAlaIleAlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThr 360
RESULT 2
US-09-976-674-1
```

```
; Sequence 1, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: QI, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 882
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-976-674-1
Alignment Scores:
Pred. No.: 1,369-208 Length: 882
Score: 1836.50 Matches: 360
Percent Similarity: 67.5% Conservative: 0
Best Local Similarity: 67.5% Mismatches: 0
Query Match: 94.8% Indels: 173
DB: 2 Gaps: 1
US-10-825-632-8 (1-1083) x US-09-976-674-1 (1-882)
Qy 2 GAAGAAGATCCAGATCAGCTGAGTCCGCTACCTTTGTTCTCCAAAGAAGATTGTAGTA 61
Db 244 GluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluPheAspArg 263
Qy 62 TATTCTGGCTATTGGTGTGTCCTCAAAAGCTGAAACAACTCCAGTGGTGTGAATTCCT 121
Db 264 TyrSerGlyTyrTrpCysProLysAlaGluThrThrProSerGlyGlyLysLeu 283
Qy 122 AGAATCTATATGAAGAAATGATGATCTGAGTGGAAATATTCATGTTACATCCCT 181
Db 284 ArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleHisValThrSerPro 303
Qy 182 ATGTTGGAAACAGGAGGCGAGATTCATTCGTTATCTCTAAACAGGTACAGCAATCCT 241
Db 304 MetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThrAlaAsnPro 323
Qy 242 AAAGTCACCTTTTAAGATGTCAGAAATAATGATGATTCATCTCAAGGAGGATCATA---- 295
Db 324 LysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIleIleAspVal 343
Qy 295 ----- 295
Db 344 IleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGluTyrIleAla 363
Qy 295 ----- 295
Db 364 ArgAlaGlyTyrThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAspArgSerGln 383
Qy 295 ----- 295
Db 384 ThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGluAspAspVal 403
Qy 295 ----- 295
Db 404 MetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeuIleIleTyr 423
Qy 295 ----- 295
Db 424 GluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPheProGlnSer 443
Qy 295 ----- 295
```

444 HisGluGluGluLeuPheIlePheAlaSerGluCysLysThrGlyPheArgHisLeu 463
295 -----
464 TyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGlyGlyLeuPro 483
295 -----
484 AlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSerGlyGluTrp 503
296 -----GTTGATGAAGTCAGAGCGCTGATAT 322
504 GluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgLeuValTyr 523
323 TTTCAGGACCAACAGCTCCCTTTAGAGCATACCTGTAGCTAGTACGTTACGTTAAAT 382
524 PheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSerTyrValAsn 543
383 CCTGGAGAGTGACAAGCGCTGACTGACCTGGCTACTCACATTTCTGCTGCATCAGTCAG 442
544 ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCysIleSerGln 563
443 CACTGTCACTCTTTATAGTAGTATAGTACCAAGAGAAATCCACACTGTGTCTCCCTT 502
564 HisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeu 583
503 TACAAGCTATCAAGTCTCTGAAGATGACCAACTTGCACAAACAAAGAAATTTGGCCACC 562
584 TyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPheTrpAlaThr 603
563 ATTTTGATTCAGAGCTCTCTCTGACTATACCTCCAGAAATTTCTCTTTTGAA 622
604 IleLeuAspSerAlaGlyProLeuProAspTyrThrProGluLysPheSerPheGlu 623
623 AGTACTAGGATTTACATTTGATGGATGCTCTACAAGCCTCATGATCTACAGCCTCGA 682
624 SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProgly 643
683 AAGAAATATCTACTGCTGTTTCATATATGTTGGTCTCCTCAGTGCAGTGTGTAATAT 742
644 LysLysTyrProThrValLeuPheIleTyrGlyProGlnValGlnLeuValAsnAsn 663
743 CGTTTAAAGAGTCAGTATTTCCGCTTGAATACCTAGCTCTCTAGGTTTATGTTT 802
664 ArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrValVal 683
803 GTAGTGATAGACAACAGGGGATCCTGTACCGAGGGCTTAAATTTGAAGGCCCTTTAAA 862
684 ValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLys 703
863 TATAAATGGGTCAATAGAAATTCAGCATCAGTGGAGGACTCCCAATATCTAGCTTCT 922
704 TyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyrLeuAlaSer 723
923 CGATATGATTTCAATGACTAGATCGTGTGGGCATCCACGCTGCTGCTTATGGAGATAC 982
724 ArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrSerTyrGlyGlyTyr 743
983 CTCCTCCGATGGCAATTAATGACAGAGTCAATATCTCAGGTTGCTATTTGCTGGGGCC 1042
744 LeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAlaGlyAla 763
1043 CCAGTCACCTCTGGATCTTCTATCATACAGATACACAG 1081
764 ProValThrLeuTrpIlePheTyrAspThrGlyTyrThr 776

RESULT 3
US-10-070-464-1
; Sequence 1, Application US/10070464
; Patent No. 6881564
; GENERAL INFORMATION:
; APPLICANT: ABBOTT, Catherine Anne
; APPLICANT: GORRELL, Mark Douglas

TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES
FILE REFERENCES: GH-007
CURRENT APPLICATION NUMBER: US/10/070,464
CURRENT FILING DATE: 2002-03-07
PRIOR APPLICATION NUMBER: PCT/AU00/01085
PRIOR FILING DATE: 2000-09-11
PRIOR APPLICATION NUMBER: AU PQ5709
PRIOR FILING DATE: 2000-02-18
PRIOR APPLICATION NUMBER: AU PQ2762
NUMBER OF SEQ ID NOS: 8
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 882
TYPE: PRT
ORGANISM: Homo Sapiens
US-10-070-464-1

Alignment Scores:
Pred. No.: 1.36e-208 Length: 882
Score: 1836.50 Matches: 360
Percent Similarity: 67.5% Conservative: 0
Best Local Similarity: 67.5% Mismatches: 0
Query Match: 94.8% Indels: 173
DB: 2 Gaps: 1

US-10-825-632-8 (1-1083) x US-10-070-464-1 (1-882)

QY 2 GNAGAGATGCCAGATCAGCTGAGTGCCTACCTTTGTTCTCCAAGAAGAAATTTGTAGTA 61
Db 244 GluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluPheAspArg 263
QY 62 TATTCTGCTATTGGTGTGTCCTCAAAAGCTGAAACAACTCCAGTGGTGGTAAATTCCT 121
Db 264 TyrSerGlyTyrTrpTyrCysProLysAlaGluThrThrProSerGlyGlyLysIleLeu 283
QY 122 AGAATTCATATGAAGAATAATGATGATCTGAGTGGAAATTTTTCATGTTTACATCCCT 181
Db 284 ArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisValThrSerPro 303
QY 182 ATGTTGGAAACAGGAGGCGAGATTCATTCGTTTATCTTAAACAGGTACAGCAATCCT 241
Db 304 MetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThrAlaAsnPro 323
QY 242 AAAGTCATTTTAAAGATGTCGAAATAATGATGATCTGCTGAGGAGGAGGATCATA---- 295
Db 324 LysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIleIleAspVal 343
QY 295 ----- 295
Db 344 IleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGluTyrIleAla 363
QY 295 ----- 295
Db 364 ArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAspArgSerGln 383
QY 295 ----- 295
Db 384 ThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGluAspAspVal 403
QY 295 ----- 295
Db 404 MetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeuIleIleTyr 423
QY 295 ----- 295
Db 424 GluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPheProGlnSer 443
QY 295 ----- 295
Db 444 HisGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPheArgHisLeu 463
QY 295 ----- 295

Db 464 TyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGlyLeuPro 483
Qy 295 -----
Db 484 AlaProSerAspPheLysCysProIleLysGluIleAlaIleThrSerGlyGluTrp 503
Qy 296 -----GTTGATGAAGTCAGAAAGCTGGTATAT 322
Db 504 GluValLeuGlyArgHisGlySerAsnIleGlnValAlaAspGluValArgLeuValTyr 523
Qy 323 TTGAAGGACCAAGACATCCCTTTAGAGCATCACCTGTACGTAGTACGTACGTAAAT 382
Db 524 PheGluGlyThrLysAspSerProLeuGluHisLysLeuTyrValValSerTyrValAsn 543
Qy 383 CTGGAGAGGTCAAGGCTGACTGACCGTGGCTACTCACATTTCTTGGCTCATGCTAG 442
Db 544 ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysLysIleSerGln 563
Qy 443 CACTGTGACTTCTTTATAGTAAGTATAGTAACAGAGAAATCCACACTGTGTGCTT 502
Db 564 HisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeu 583
Qy 503 TACAAGCTATCAAGTCTGAAGATGACCAACTTGCACAAACAAAGCAATTTTGGCCACC 562
Db 584 TyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPheTrpAlaThr 603
Qy 563 ATTTTGGATTGACAGAGTCTCTTCTGACTATATCTCTCCAGAAATTTCTTTTGA 622
Db 604 IleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePheSerPheGlu 623
Qy 623 AGTACTACTGATTTACATTTGATGGATGCTCTACAGCCTCATGATCTACAGCTGGA 682
Db 624 SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly 643
Qy 683 AAGAAATATCTACTGCTGCTTCATATATGATGGTCTCTAGTGCCTCTCTAGGTTATGTGTT 742
Db 644 LysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeuValAsnAsn 663
Qy 743 CGTTTAAAGAGCTCAAGTATTTCCCTTGAATACCTTAGCTCTCTAGGTTATGTGTT 802
Db 664 ArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrValVal 683
Qy 803 GTAGTGTATAGACACAGGGGATCTGTCCAGGAGCTTAAATTTGAAGGCGCTTTAAA 862
Db 684 ValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLys 703
Qy 863 TATAAATGGTCAAAATAGAAATTCAGCATCAGTGGAGGACTCCCAATATCTAGCTTCT 922
Db 704 TyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyrLeuAlaSer 723
Qy 923 CGATATGATTTCAATTCATTTAGATCGTGTGGGCATCCAGCGTGGTCTCTATGAGGATAC 982
Db 724 ArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrSerTyrGlyTyr 743
Qy 983 CTCTCCCTGATGGCATTATGAGAGGTACAGATATCTTACGGTTGCTATGCTGGGCC 1042
Db 744 LeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAlaGlyAla 763
Qy 1043 CCAGTCACTCTGTGATCTTCTATGATACAGGATACACG 1081
Db 764 ProValThrLeuTrpIlePheTyrAspThrGlyTyrThr 776

RESULT 4

US-09-976-674-7
; Sequence 7, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV

FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 690
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-7
Alignment Scores:
Pred. No.: 3,46e-136 Length: 690
Score: 1226.00 Matches: 259
Percent Similarity: 48.5% Conservative: 0
Best Local Similarity: 48.5% Mismatches: 0
Query Match: 63.3% Indels: 275
DB: 2 Gaps: 2
US-10-825-632-8 (1-1083) x US-09-976-674-7 (1-690)
Qy 2 GAAGAAGATGCCAGATCAGCTGAGTGGCTACTTCTTCTTCTCAAGAGAAATTTGTAGA 61
Db 244 GluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluPheAspArg 263
Qy 62 TATTCCTGCTATTGGTGTGTCGAAAGCTGAAACAACTCCCAAGTGGTGGTAAATTTCTT 121
Db 264 TyrSerGlyTyrTrpCysProLysAlaGluThrThrProSerGlyGlyLysIleLeu 283
Qy 122 AGAATTCCTATGAAGAAATGATGAATCTGAGTGGAAATTTATTCATGTTACATCCCT 181
Db 284 ArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleHisValThrSerPro 303
Qy 182 ATGTTGGAAACAGAGGGCAGATTCATTCCTGTTATCTTAAACACAGGTACAGCAATCT 241
Db 304 MetLeuGlnThrArgArgAlaAspSerPheArgTyrProLysThrGlyThrAlaAsnPro 323
Qy 242 AAATCTCTTTAAGATGTCAGAAATTAATGATTGATGCTGAAGGAGGATCATA----- 295
Db 324 LysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIleIleAspVal 343
Qy 295 -----
Db 344 IleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGluTyrIleAla 363
Qy 295 -----
Db 364 ArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAspArgSerGln 383
Qy 295 -----
Db 384 ThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGluAspAspVal 403
Qy 295 -----
Db 404 MetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeuIleIleTyr 423
Qy 295 -----
Db 424 GluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPheProGlnSer 443
Qy 295 -----
Db 444 HisGluGluGluIleGluPheIlePheAlaSerCysLysThrGlyPheArgHisLeu 463
Qy 295 -----
Db 464 TyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGlyLeuPro 483
Qy 295 -----

484 AlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSerGlyGluTrp 503
296 -----GTTGATGAAGTCAGAGCGCTGGTATAT 322
504 GluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArgLeuValTyr 523
323 TTTGAAGGCAACCAAGACTCCCTTTAGAGCATCAGCTAGTACGTTAGTCAGTAAAT 382
524 PheGluGlyThrLysAspSerProLeuGluHisLeuValValSerTyrValAsn 543
383 CTTGGAGGTGACCAAGGCTGACTGACCGTGGTACTCAGTATCTTGTGTCATCAGTCAG 442
544 ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysIleSerGln 563
443 CACTGTGACTCTTTTAAAGTAAAGTAAACCAAGAGATCCACACTGTGTCTCCTT 502
564 HisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeu 583
503 TACAGCTATCAAGTCTGAGATGACCAACTTCGCAAAACAAAGGAATTTTGGCCACC 562
584 TyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPheTrpAlaThr 603
563 ATTTTGGATTACAGTCTCTCTCTGACTATCTCTCCAGAAATTTCTCTTTGAA 622
604 IleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePheSerPheGlu 623
623 AGTACTACTGATTACATTTGATGGATGCTCTACAGCCTCATGATCAGCCTGGA 682
624 SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly 643
683 AAGAAATCTCTACTGCTGCTCATATATGTTGGTCTCAGTGCAGTGGTGAATAAT 742
644 LysLysTyrProThrValLeuPheIleTyrGlyGly ----- 655
743 CGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTAGCCTCTCTAGTTTATGTGTT 802
655 ----- 655
803 GTAGTAGACAAACAGGGGATCTGTACCGAGGGCTTAATTTGAAGGCCCTTTAA 862
655 ----- 655
863 TATAAATGGGTCAAATAGAAATTGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCT 922
655 ----- 655
923 CGATATGATTTCAATTGACTTAGATCGTGTGGGCATCCACGGCTGGTCTATGAGGATAC 982
655 ----- 655
983 CTCTCCCTGATGGCATTAAATGACAGGTCAGATATCTCAGGCTTGTATGCTGGGCC 1042
656 -----Arg-LeuLeuLeuLeuGlyPr 662
1043 CCAGTCACTCTGATCTTCTATGATACAGGATACACGG 1082
662 oGlnSerLeuCysGlySerSerMetIleGlnAspThrArg 675

RESULT 5
US-09-976-674-3
; Sequence 3, Application US/09976674
; Patent No. 684180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12

; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 863
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-3
Alignment Scores:
Pred. No.: 1,73e-135 Length: 863
Score: 1220.50 Matches: 236
Percent Similarity: 54.3% Conservative: 54
Best Local Similarity: 44.2% Mismatches: 69
Query Match: 63.0% Indels: 175
DB: 2 Gaps: 2
US-10-825-632-8 (1-1083) x US-09-976-674-3 (1-863)
QY 5 GAAGATGCCAGATCAGCTGAGTCGCTTGTCTCCAAAGAAATTTGATAGATAT 64
Db 234 AspAspProLysSerAlaGlyValAlaThrPheValIleGlnGluPheAspArgPhe 253
QY 65 TCTGCTATTGCTGCTGCTCCAAAGCTGAAACCAACTCCCAGTGGTGGT --- AAAATTCTT 121
Db 254 ThrGlyTyrTrpTrpCysProThrAlaSerTrpGluGlySerGluGlyLeuLysThrLeu 273
QY 122 AGAATTTCTATGATGAAGAAATGATGAATCTGAGGTGGAATATTATTATTTACATCCCT 181
Db 274 ArgIleLeuTyrGluGluValAspGluSerGluValGluValIleHisValProSerPro 293
QY 182 ATGTTGGAAACACAGGAGGCGCAGATTCTCGTTTATCTTAAACACAGGTACAGCAATCCT 241
Db 294 AlaLeuGluGluArgLysThrAspSerTyrArgTyrProArgThrGlySerLysAsnPro 313
QY 242 AAAGTCACCTTTTAAGATCTCAGAAATAATGATTCATCTGAAAGAGGATCATA --- 295
Db 314 LysIleAlaLeuLysLeuAlaGluPheGlnThrAspSerGlnGlyLysIleValSerThr 333
QY 295 ----- 295
Db 334 GlnGluLysGluLeuValGlnProPheSerSerLeuPheProLysValGluTyrIleAla 353
QY 295 ----- 295
Db 354 ArgAlaGlyTrpThrArgAspGlyLysTyrAlaTrpAlaMetPheLeuAspArgProGln 373
QY 295 ----- 295
Db 374 GlnTrpLeuGlnLeuValLeuLeuProAlaLeuPheIleProSerThrGluAsnGlu 393
QY 295 ----- 295
Db 394 GluGlnArgLeuAlaSerAlaArgAlaValProArgAsnValGlnProTyrValValTyr 413
QY 295 ----- 295
Db 414 GluGluValThrAsnValTrpIleAsnValHisAspIlePheTyrProPheProGlnSer 433
QY 295 ----- 295
Db 434 GluGlyGluAspGluLeuCysPheLeuArgAlaAsnGluCysLysThrGlyPheCysHis 453
QY 295 ----- 295
Db 454 LeuTyrLysValThrAlaValLeuLysSerGlnGlyTyrAspTrpSerGluProPheSer 473
QY 295 ----- 295
Db 474 ProGlyGluAspGluPheLysCysProIleLysGluGluIleAlaLeuThrSerGlyGlu 493
QY 296 -----GTTGATGAAGTCAGAGGCTGTA 319
Db 494 TrpGluValLeuAlaArgHisGlySerLysIleTrpValAsnGluGluThrLysLeuVal 513

QY	320	TATTTTGAAGCACC	AAGA	CTCCCTTTTAGAGCATCACCTGTACCTAGTCAGTTACGTA	379
DB			:		
QY	514	TyrPheGlnGlyThrLysAspThrProLeuGluHisLeuTyValValSerTyrGlu	533		
DB			:		
QY	380	AATCCTGGAGAGTGACAGCGTGCATGCCCTGTACTACATCTTCTGTCGATCAGT	439		
DB			:		
QY	534	AlaIalaGlyLeuValArgLeuThrThrProGlyPheSerHisSerCySerMetSer	553		
DB			:		
QY	440	CAGCACTGTGCATCTTTATTAGTAAAGTATAGTAAACCAGAAAGATCCACACTGTGTGTCC	499		
DB			:		
QY	554	GlnAenPheAspMetPheValSerHisTyrSerSerValSerThrProProCySValHis	573		
DB			:		
QY	500	CTTTACAAGCTATCAAGTCTCCAAGATGACCCMACTGCGACAAACAAGGAATTTTGGGCC	559		
DB			:		
QY	574	ValTyrLysLeuSerGlyProAspAspAspProLeuHisLysGlnProAspPheTyrPala	593		
DB			:		
QY	560	ACCAATTTGGATTACAGAGGCTCTCTCTGCATATACTCTCCAGAAAATTTCTCTTTT	619		
DB			:		
QY	594	SerMetMetGluAlaAlaSerCySerProAspTyrValProProGluLillePheHisPhe	613		
DB			:		
QY	620	GAAAGTACTACTGGATTACATTGATCGGGATGCTTACAAGGCTCATGATCTACAGACT	679		
DB			:		
QY	614	HisThrArgSerAspValArgLeuTyrGlyMetLileTyrLysProHisAlaLeuGlnPro	633		
DB			:		
QY	680	GGAAGAAATATCTACTGTGCTGTTCATATATCGTGTGCTCCTCAGGTGCAGTGTGGTGAAT	739		
DB			:		
QY	634	GlyLysLysHisProThrValLeuPheValTyrGlyGlyProGlnValGlnLeuValAen	653		
DB			:		
QY	740	NATCGGTTTAAAGGATCAAGTATTTCGGTTGAATACCTTCAAGCTCTCTAGGTATTGTG	799		
DB			:		
QY	654	AntSerPheLysGlyLileTyrLysLeuArgLeuAenThrLeuAlaSerLeuGlyTyrZala	673		
DB			:		
QY	800	GTTGTAGTGATGACACAGGGATCTGTCCACGGGGCTTAAATTTGAAGCGCCCTTT	859		
DB			:		
QY	674	ValValValLeuAspArgGlySerCySelnArgGlyLeuArgPheGluGlyAlaLeu	693		
DB			:		
QY	860	AAATATAAANTGGTCAAAATGAANATTGACGATCAGGTGGAAGGACTCCAAATCTAGCT	919		
DB			:		
QY	694	LysAenGlnMetGlyGlnValGluLilleGluaspGlnValGluGlyLeuGlnPheValAa	713		
DB			:		
QY	920	TCTCGATATGATTTTCATTGACTTAGATCGTGTGGGATCCACGGCTGTCTCTATGGAGGA	979		
DB			:		
QY	714	GluLysTyrGlyPheLeuAspLeuSerArgValAlaIleHisGlyTyrPsrTyrGlyLy	733		
DB			:		
QY	980	TACTCTCTCCCTGATGGCAATTAATGCAGAGCTCAGATATCTTCAGGGTGTCTATTCCTGGG	1039		
DB			:		
QY	734	PheLeuSerLeuMetGlyLeuLileHisLysProGlnValPheLysValAlaIleGlyLy	753		
DB			:		
QY	1040	GCCCGACTCACTGTGGATCTCTCTATGATACAGGATACACG	1081		
DB			:		
QY	754	AlaProValThrValTyrMetalatTyrAspThrGlyTyrThr	767		
DB			:		

RESULT 6

```

US-09-976-674-23
; Sequence 23, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Q1; Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPP1V
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 23
; LENGTH: 892
; TYPE: PPT

```

i ORGANISM: Homo sapiens
US-09-976-674-23

Alignment Scores:
 Prad. No.: 1,76e-135 Length: 892
 Score: 1229.50 Matches: 236
 Percent Similarity: 54.2% Conservative: 236
 Best Local Similarity: 44.2% Mismatches: 69
 Query Match: 63.0% Indels: 175
 DB: 2 Gaps: 2

US-10-825-632-8 (1-1083) x US-09-976-674-23 (1-892)

QY	5	GAAGATCCAGATCAGCTGAGTCGCTTGTCTTCTCCAAAGAGATTGTAGATAT	64
DB	263	AspAspProLysSerAlaGlyValAlaThrPheValIleGlnGluPheAspArgPhe	282
QY	65	TCTGGCTATTGGTGGTGTCCAAAAGCTGAACCACTCCCACTGGTGT	121
DB	283	ThrGlyTyrTrpTrpCysProThrAlaSerTrpGluGlySerGluGlyLeuLysThrLeu	302
QY	122	AGAATTCATATCAAGAAATGATGAATCTGAGTGGAAATATTATGTTACATCCCT	181
DB	303	ArgIleLeuTyrGlnGluValAspGluSerGluValGluValIleHisValProSerPro	322
QY	182	ATGTTGGAACAAGGAGGCAGATTCATTCCTCGTTATCTTAAACAGGTCACGCAATCT	241
DB	323	AlaLeuGlnGluArgLysThrAspSerTyrArgTyrProArgTyrGlySerLysAsnPro	342
QY	242	AAAGTCACITTTAAGATGTCAGAAATAATCATTCATGCTGAAGAAAGGATCAT	295
DB	343	LysIleAlaLeuLysLeuAlaGluPheGlnThrAspSerGlnGlyLysIleValSerThr	362
QY	295	-----	295
DB	363	GlnGluLysGluLeuValGlnProPhePheSerSerLeuPheProLysValGluTyrIleAla	382
QY	295	-----	295
DB	383	ArgAlaGlyTrpThrArgAspGlyLysTyrAlaTrpAlaMetPheLeuAspArgProGln	402
QY	295	-----	295
DB	403	GlnTrpLeuGlnLeuValLeuLeuProProAlaLeuPheIleProSerThrGluAsnGlu	422
QY	295	-----	295
DB	423	GluGlnArgLeuAlaSerAlaArgAlaValProArgAsnValGlnProTyrValValTyr	442
QY	295	-----	295
DB	443	GluGluValTyrAsnValTrpIleAsnValHisAspIlePheTyrProPheProGlnSer	462
QY	295	-----	295
DB	463	GluGlyGluAspGluLeuCysPheLeuArgAlaAsnGluCysLysThrGlyPheCysHis	482
QY	295	-----	295
DB	483	LeuTyrLysValThrAlaValLeuLysSerGlnGlyTyrAspTrpSerGluProPheSer	502
QY	295	-----	295
DB	503	ProGlyGluAspGluPheLysCysProIleLysGluGluIleAlaLeuThrSerGlyGlu	522
QY	296	-----GTTGATGAAGTCAGAGCTGGTA	319
DB	523	TrpGluValLeuAlaArgHisGlySerLysIleTrpValAsnGluGluThrLysLeuVal	542
QY	320	TATTTGAAGGCCAAAGACTCCCTTTAGAGCATCACCTGTACGTAGTCAGTTACGTA	379
DB	543	TyrPheGlnGlyThrLysAspThrProLeuGluHisHisLeuTyrValValSerTyrGlu	562

380	AAATCTCGAGAGGTGACAAAGGCTGACGTACCGTGGCTACTCATTCTTCTGTCATCAGT	439	Score: 1220.50	Matches: 236			
563	AlaAlaGlyGluIleValArgLeuThrThrProGlyPheSerHisSerCysSerMetSer	582	Percent Similarity: 54.3%	Conservative: 54			
440	CAGCACTGTGACTCTCTTATAAGTAACTAGTAACACAGAGAATCCACACTGTGTGCC	499	Best Local Similarity: 44.2%	Mismatches: 69			
583	GlnAsnPheAspMetPheValSerHisTyrSerValSerThrProCysValHis	602	Query Match: 63.0%	Indels: 175			
500	CTTTACAGCTATCAAGTCTGGAAGATGACCAACTTGCAAAACAAGAAATTTGGGCC	559	DB: 2	Gaps: 2			
603	ValTyrLysLeuSerGlyProAspAspProLeuHisLysGlnProArgPheTrpAla	622	US-10-825-632-8 (1-1083) x US-09-976-674-27 (1-892)				
560	ACCATTTTGGATTCAGCAGGTCCTCTTCTGACTATACCTCTCCAGAAATTTCTCTTT	619	QY	5	GAAGATGCCAGATCAGCTGGAGTCGTACCTTTGTTCTCCAAAGAATTTGTAGATAT	64	
623	SerMetMetGluAlaAlaSerCysProAspTyrValProProGluIlePheHisPhe	642	Db	263	AspAspProLysSerAlaGlyValAlaThrPheValIleGlnGluPheAspArgPhe	282	
620	GAAATACTACTGGATTTACATTTGATGGATGCTCTACAGCCTCATGATCTACAGCT	679	QY	65	TCTGCTATTGCTGTGTCCAAAAGCTCAAACTCCCACTCCCTGCTGCT	121	
643	HisThrArgSerAspValArgLeuTyrGlyMetIleTyrLysProHisAlaLeuGlnPro	662	Db	283	ThrGlyTyrTrpTrpCysProThrAlaSerTrpGluGlySerGluGlyLeuThrLeu	302	
680	GGAAAGAATATCTACTGCTGCTTCATATATATGCTGCTCAGGTGAGTTGGTGAAT	739	QY	122	AGAATCTATATGAAGAAATGATGATCTGAGGTGAAATTTATTTATCTATCCCT	181	
663	GlyLysLysHisProThrValLeuPheValTyrGlyGlyProGlnValGlnLeuValAsn	682	Db	303	ArgIleLeuTyrGluGluValAspGluSerGluValGluValIleHisValProSerPro	322	
740	AATCGGTTTAAAGGAGTCAAGTATTTCGGCTTGAATACCTCCTAGCTCTCTAGGTTATGT	799	QY	182	ATGTTGGAACAAGAGGAGGAGATTCCTCGTTTATCTTAAACAGGTACAGCAAACTCT	241	
683	AsnSerPheLysGlyIleLysTyrLeuArgLeuAsnThrLeuAlaSerLeuGlyTyrAla	702	Db	323	AlaLeuGluGluArgLysThrAspSerTyrArgTyrProArgThrGlySerLysAsnPro	342	
800	GTGTAGTAGTACAGACACAGGGGATCCTGTGTCACGAGGGCTTAAATTTGAAGGCGCTTT	859	QY	242	AAAGTCACTTTTAAAGATGTGCAAAATATGATGATCTGAAAGAGGATCATA	295	
703	ValValValIleAspGlyArgGlySerCysGlnArgGlyLeuArgPheGluGlyAlaLeu	722	Db	343	LysIleAlaLeuLysLeuAlaGluPheGlnThrAspSerGlnGlyLysIleValSerThr	362	
860	AAATATAAATGGTCAATAGAAATTCAGATCAGGTGGAAGGACTCCAATATCTAGCT	919	QY	295	-----	295	
723	LysAsnGlnMetGlyGlnValGluIleGluAspGlnValGluGlyLeuGlnPheValAla	742	Db	383	ArgAlaGlyTrpThrArgAspGlyLysTyrAlaTrpAlaMetPheLeuAspArgProGln	402	
920	TCTCGATATGATTTCAATTCATAGATCGTGTGGGATCCACGCTGCTCTATGGAGGA	979	QY	295	-----	295	
743	GluLysTyrGlyPheIleAspLeuSerArgValAlaIleHisGlyTrpSerTyrGlyGly	762	Db	403	GlnTrpLeuGlnLeuValLeuLeuProProAlaLeuPheIleProSerThrGluAsnGlu	422	
980	TACCTCTCCTGATGCGATTAATGACAGAGGTACAGATATCTTCAGGGTGTCTATGCTGG	1039	QY	295	-----	295	
763	PheLeuSerLeuMetGlyLeuIleHisLysProGlnValPheLysValAlaIleAlaGly	782	Db	423	GluGlnArgLeuAlaSerAlaArgAlaValProArgAsnValGlnProTyrValValTyr	442	
1040	GCCCCAGTCACTCTGTGGATCTTCTATGATACAGATACAG 1081		QY	295	-----	295	
783	AlaProValThrValTrpMetAlaTyrAspThrGlyTyrThr 796		Db	443	GluGluValThrAsnValTrpIleAsnValHisAspIlePheTyrProPheProGlnSer	462	
RESULT 7				QY	295	-----	295
US-09-976-674-27				Db	463	GluGlyGluAspGluLeuCysPheLeuArgAlaAsnGluCysLysThrGlyPheCysHis	482
; Sequence 27, Application US/09976674				QY	295	-----	295
; Patent No. 6844180				Db	483	LeuTyrLysValThrAlaValLeuLysSerGlnGlyTyrAspTrpSerGluProPheSer	502
; GENERAL INFORMATION:				QY	295	-----	295
; APPLICANT: Qi, Steve				Db	503	ProGlyGluAspGluPheLysCysProIleLysGluGluIleAlaLeuThrSerGlyGlu	522
; APPLICANT: Akinsanya, Karen				QY	296	-----	319
; APPLICANT: Riviere, Pierre				Db	523	TrpGluValLeuAlaArgHisGlySerLysIleTrpValAsnGluGluThrLysLeuVal	542
; APPLICANT: Junien, Jean-Louis				QY	320	TATTTGAAGCCACCAAGACTCCCTTTAGACATCCTCTAGCTAGTACGTTACGTA	379
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV				Db	543	TyrPheGlnGlyThrLysAspTrpProLeuGluHisLeuTyrValValSerTyrGlu	562
; FILE REFERENCE: 70669				QY	380	AATCTGGAGAGGTGACAAAGCTGAGTCCGCGGCTACTCATTCTTCTGTCATCAGT	439
; CURRENT APPLICATION NUMBER: US/09/976,674				Db	563	AlaAlaGlyGluIleValArgLeuThrThrProGlyPheSerHisSerCysSerMetSer	582
; CURRENT FILING DATE: 2001-10-12				QY	440	CAGCACTGTGACTCTCTTATAAGTAACTAGTAACACAGAGAATCCACACTGTGTGCC	499
; PRIOR APPLICATION NUMBER: US 60/240,117				Db	583	GlnAsnPheAspMetPheValSerHisTyrSerValSerThrProCysValHis	602
; PRIOR FILING DATE: 2000-10-12				QY	500	CTTTACAGCTATCAAGTCTGGAAGATGACCAACTTGCAAAACAAGAAATTTGGGCC	559
; NUMBER OF SEQ ID NOS: 61				Db	603	ValTyrLysLeuSerGlyProAspAspProLeuHisLysGlnProArgPheTrpAla	622
; SOFTWARE: PatentIn version 3.1				QY	560	ACCATTTTGGATTCAGCAGGTCCTCTTCTGACTATACCTCTCCAGAAATTTCTCTTT	619
; SEQ ID NO 27				Db	623	SerMetMetGluAlaAlaSerCysProAspTyrValProProGluIlePheHisPhe	642
; LENGTH: 892				QY	620	GAAATACTACTGGATTTACATTTGATGGATGCTCTACAGCCTCATGATCTACAGCT	679
; TYPE: PRT				Db	643	HisThrArgSerAspValArgLeuTyrGlyMetIleTyrLysProHisAlaLeuGlnPro	662
; ORGANISM: Homo sapiens				QY	680	GGAAAGAATATCTACTGCTGCTTCATATATATGCTGCTCAGGTGAGTTGGTGAAT	739
US-09-976-674-27				Db	663	GlyLysLysHisProThrValLeuPheValTyrGlyGlyProGlnValGlnLeuValAsn	682
Alignment Scores:				QY	740	AATCGGTTTAAAGGAGTCAAGTATTTCGGCTTGAATACCTCCTAGCTCTCTAGGTTATGT	799
Pred. No.:				Db	683	AsnSerPheLysGlyIleLysTyrLeuArgLeuAsnThrLeuAlaSerLeuGlyTyrAla	702
				QY	800	GTGTAGTAGTACAGACACAGGGGATCCTGTGTCACGAGGGCTTAAATTTGAAGGCGCTTT	859
				Db	703	ValValValIleAspGlyArgGlySerCysGlnArgGlyLeuArgPheGluGlyAlaLeu	722
				QY	860	AAATATAAATGGTCAATAGAAATTCAGATCAGGTGGAAGGACTCCAATATCTAGCT	919
				Db	723	LysAsnGlnMetGlyGlnValGluIleGluAspGlnValGluGlyLeuGlnPheValAla	742
				QY	920	TCTCGATATGATTTCAATTCATAGATCGTGTGGGATCCACGCTGCTCTATGGAGGA	979
				Db	743	GluLysTyrGlyPheIleAspLeuSerArgValAlaIleHisGlyTrpSerTyrGlyGly	762
				QY	980	TACCTCTCCTGATGCGATTAATGACAGAGGTACAGATATCTTCAGGGTGTCTATGCTGG	1039
				Db	763	PheLeuSerLeuMetGlyLeuIleHisLysProGlnValPheLysValAlaIleAlaGly	782
				QY	1040	GCCCCAGTCACTCTGTGGATCTTCTATGATACAGATACAG 1081	
				Db	783	AlaProValThrValTrpMetAlaTyrAspThrGlyTyrThr 796	
				US-09-976-674-27			
				; Sequence 27, Application US/09976674			
				; Patent No. 6844180			
				; GENERAL INFORMATION:			
				; APPLICANT: Qi, Steve			
				; APPLICANT: Akinsanya, Karen			
				; APPLICANT: Riviere, Pierre			
				; APPLICANT: Junien, Jean-Louis			
				; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV			
				; FILE REFERENCE: 70669			
				; CURRENT APPLICATION NUMBER: US/09/976,674			
				; CURRENT FILING DATE: 2001-10-12			
				; PRIOR APPLICATION NUMBER: US 60/240,117			
				; PRIOR FILING DATE: 2000-10-12			
				; NUMBER OF SEQ ID NOS: 61			
				; SOFTWARE: PatentIn version 3.1			
				; SEQ ID NO 27			
				; LENGTH: 892			
				; TYPE: PRT			
				; ORGANISM: Homo sapiens			
				US-09-976-674-27			
				Alignment Scores:		1.76e-135	
				Pred. No.:		892	

```
Db 583 GlnAsnPheAspMetPheValSerHisTyrSerSerValSerThrProCysValHis 602
QY 500 CTTTACAGCTATCAAGTCTGAGATGACCACTGCAAAACAAAGGAATTTGGGCC 559
Db 603 ValTyrLysLeuSerGlyProAspAspAspProLeuHisLysGlnProArgPheThrPala 622
QY 560 ACCATTGATTCACAGGCTCTCTTCTGCTGACTATCTCTCCAGAAATTTCTCTTT 619
Db 623 SerMetMetGluAlaAlaSerCysProProAspTyrValProProGluLeuPheHisPhe 642
QY 620 GAAAGTACTACTGATTTACATTTGATGGATGCTCTCAAGCCTCATGATCTACGCCT 679
Db 643 HisThrArgSerAspValArgLeuTyrGlyMetIleTyrLysProHisAlaLeuGlnPro 662
QY 680 GGAAGAAATATCTACTGCTGCTTCAATATGATGCTCTCAGGTGCAAGTGGTGAAT 739
Db 663 GlyLysLysHisProThrValLeuPheValTyrGlyGlyProGlnValGlnLeuValAsn 682
QY 740 AATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTTAGCCTCTCTAGGTATGTG 799
Db 683 AsnSerPheLysGlyLysTyrLeuArgLeuAsnThrLeuAlaSerLeuGlyTyrAla 702
QY 800 GTTGTAGTGATGACACACAGGGATCCTGTCCAGGAGGCTTAAATTTGAAGCGCCTTT 859
Db 703 ValValValIleAspGlyArgGlySerCysGlnArgGlyLeuArgPheGluGlyAlaLeu 722
QY 860 AATATAAATGGTCAATAGAAATTCACGATCAGGTGGAAGGACTCCAAATATCTAGCT 919
Db 723 LysAsnGlnMetGlyGlnValGluIleGluAspGlnValGluGlyLeuGlnPheValAla 742
QY 920 TCTCGATATGATTTCAATGACTTATGATGCTGGGATCCAGGCTGCTCTATGGAGGA 979
Db 743 GluLysTyrGlyPheIleAspLeuSerArgValAlaIleHisGlyTyrSerTyrGlyGly 762
QY 980 TACTCTCTCCTGATGGCAATTAATCAGAGTCAGATATCTCAGGCTGCTCTATGGTGG 1039
Db 763 PheLeuSerLeuMetGlyLeuIleHisLysProGlnValPheLysValAlaIleAlaGly 782
QY 1040 GCCCAGTCACTCTGTGATCTTCTATGATACAGGATACAGC 1081
Db 783 AlaProValThrValTyrMetAlaTyrAspThrGlyTyrThr 796

RESULT 8
; Sequence 11, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 11
; LENGTH: 661
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-976-674-11

Alignment Scores:
Pred. No.: 3,53e-134 Length: 661
Score: 1209.00 Matches: 244
Percent Similarity: 58.2% Conservative: 0
Best Local Similarity: 58.2% Mismatches: 0
Query Match: 62.4% Indels: 175
DB: 2 Gaps: 2
```

```
US-10-825-632-8 (1-1083) x US-09-976-674-11 (1-661)
QY 2 GAAGAAGATGCCAGATCAGCTGGAGTCGGTACCTTTGTTCTCAAGAAGAATTTGATAGA 61
Db 244 GluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluPheAspArg 263
QY 62 TATTCTGGCTATTGGTGTGTCCTCAAAAGCTGAAACAACTCCCAAGTGGTGTAAATTCCTT 121
Db 264 TyrSerGlyTyrTrpCysProLysAlaGluThrThrProSerGlyGlyLysLeu 283
QY 122 AGAATTCATATGAAGAAATGATGATCTGAGGTGGAAATTTATTCATCTACATCCCT 181
Db 284 ArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleHisValThrSerPro 303
QY 182 ATGTTGAAACAAAGGAGGCGCAGATTCATTCCTGTTATCTTAAACAGGTACAGCAATCCT 241
Db 304 MetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThrAlaAsnPro 323
QY 242 AAGTCACTTTTAAGATGTCAGAAATATGATGATGCTGAAGGAGGATCATATA----- 295
Db 324 LysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIleAspVal 343
QY 295 ----- 295
Db 344 IleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGluTyrIleAla 363
QY 295 ----- 295
Db 364 ArgAlaGlyTyrThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAspArgSerGln 383
QY 295 ----- 295
Db 384 ThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGluAspAspVal 403
QY 295 ----- 295
Db 404 MetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeuIleIleTyr 423
QY 295 ----- 295
Db 424 GluGluThrThrAspIleTyrIleAsnIleHisAspIlePheHisValPheProGlnSer 443
QY 295 ----- 295
Db 444 HisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPheArgHisLeu 463
QY 295 ----- 295
Db 464 TyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGlyGlyLeuPro 483
QY 295 ----- 295
Db 484 AlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSerGlyGluTyr 503
QY 296 -----GTTGATGAAGTCAAGAGGCTGTATAT 322
Db 504 GluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArgLeuValTyr 523
QY 323 TTGAGGCACCAAGACTCCCTTTAGAGCATCACCTGTACGTAGTACGTACGTAAAT 382
Db 524 PheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSerTyrValAsn 543
QY 383 CTGAGAGGTGACAAAGCTGACTACCGTGGCTACTCATTTCTTGTGTCATCAGTCAG 442
Db 544 ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysIleSerGln 563
QY 443 CACTGTGACTCTTTTATAGTAAGTATAGTAACCAAGAAATCCACACTGTGTGCTCCCT 502
Db 564 HisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeu 583
QY 503 TACAAGCTATCAAGTCTCTGAAGATGACCCCAACTTGCAAACCAAGGAATTTGGGCCACC 562
|||||
```

Db	584	TyrLysLeuSerSerProGluAspProThrCysLysThrLysGluPheTrpAlaThr	603
Qy	563	ATTTGGATTACAGAGGCTCTTCTCGACTACTCTCCAGAAATTTCTCTTTGAA	622
Db	604	IleLeuAspSerAlaGlyProLeuProAspThrProGluIlePheSerPheGlu	623
Qy	623	AGTACTACTGGATTATGATGGATGCTCTACAAGCTCATGATCTACAGCTCGGA	682
Db	624	SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly	643
Qy	683	AAGAAATATCTACTGCTGTTTCATATATGTTGTCCTC---AGGTGCAGTTGG	734
Db	644	LysLysTyProThrValLeuPheIleTyrGlyGly-LeuLeuArgCysSerTrp	661
RESULT 9			
US-09-976-674-33			
; Sequence 33, Application US/09976674			
; Patent No. 6844180			
GENERAL INFORMATION:			
; APPLICANT: Qi, Steve			
; APPLICANT: Akinsanya, Karen			
; APPLICANT: Riviere, Pierre			
; APPLICANT: Junien, Jean-Louis			
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV			
; FILE REFERENCE: 70669			
; CURRENT APPLICATION NUMBER: US/09/976,674			
; CURRENT FILING DATE: 2001-10-12			
; PRIOR APPLICATION NUMBER: US 60/240,117			
; PRIOR FILING DATE: 2000-10-12			
; NUMBER OF SEQ ID NOS: 61			
; SOFTWARE: PatentIn version 3.1			
; SEQ ID NO 33			
; LENGTH: 879			
; TYPE: PRT			
; ORGANISM: Homo sapiens			
US-09-976-674-33			
Alignment Scores:			
Pred. No.:	3,61e-133	Length:	879
Score:	1201.00	Matches:	234
Percent Similarity:	55.1%	Conservative:	53
Best Local Similarity:	44.9%	Mismatches:	72
Query Match:	62.0%	Indels:	162
DB:	2	Gaps:	3
US-10-825-632-8 (1-1083) x US-09-976-674-33 (1-879)			
Qy	5	GAAGATGCCAGATCAGCTGGAGTCGCTACCTTTCTTCTCCAAGAAGATTTGATAGATAT	64
Db	263	AspAspProLysSerAlaGlyValAlaThrPheValIleGlnGluPheAspArgPhe	282
Qy	65	TCGTGGTATTGTGGTGTCCAAAGCTGAAACAACTCCAGTGGTGT---AAATTCCT	121
Db	283	ThrGlyTyrTrpCysProThrAlaSerTrpGluGlySerGluGlyLeuLysThrLeu	302
Qy	122	AGAATTCATATCAGAAATCATCAATCTCAGCTGGGAATTTATTCATGTTACATCCCT	181
Db	303	ArgIleLeuTyrGluGluValAspGluSerGluValGluValIleHisValProSerPro	322
Qy	182	ATCTCGAACACAGGAGGCGAGATTCATTCCGTTATCTCTAAACAGGTACAGCAATCCT	241
Db	323	AlaLeuGluGluArgLysThrAspSerTyrArgTyrProArgThrGlySerLysAsnPro	342
Qy	242	AAAGTCACTTTTAAAGATGTCAGAAATAATGATTGCTGAGGAAGGATCATAGTTGAT	301
Db	343	LysIleAlaLeuLysAlaGluPheGlnThrAspSerGlnGlyLysIleValSerThr	362
Qy	302	GAAGTCAGAGGCTG-----	316
Db	363	GlnGluLysGluLeuValGlnProPheSerSerLeuPheProLysValGluTyrIleAla	382
Qy	316	-----	316
Db	383	ArgAlaGlyTrpThrArgAspGlyLysTyrAlaTrpAlaMetPheLeuAspArgProGln	402
Qy	316	-----	316
Db	403	GlnTrpLeuGlnLeuValLeuLeuProProAlaLeuPheIleProSerThrGluAsnGlu	422
Qy	317	-----GTATAT	322
Db	423	GluGlnArgLeuAlaSerAlaArgAlaValProArgAsnValGlnProTyrValValTyr	442
Qy	323	TTTGAA-----	328
Db	443	GluGluValThrAsnValTrpIleAsnValHisAspIlePheTyrProPheProGlnSer	462
Qy	328	-----	328
Db	463	GluGlyGluAspGluLeuCysPheLeuArgAlaAsnGluCysLysThrGlyPheCysHis	482
Qy	328	-----	328
Db	483	LeuTyrLysValThrAlaValLeuLysSerGlnGlyTyrAspTrpSerGluProPheSer	502
Qy	328	-----	328
Db	503	ProGlyGluAspGluPheLysCysProIleLysGluGluIleAlaLeuThrSerGlyGlu	522
Qy	329	-----GGCACAAGACTCCCTTTAGAGCATCAC	358
Db	523	TrpGluValLeuAlaArgHisGlySerLysGlyThrLysAspThrProLeuGluHisHis	542
Qy	359	CTGTACGTAGTCAGTTAGCTAAATCTCGAGAGGTGACAAAGGTGACTCAGCTGGCTAC	418
Db	543	LeuTyrValValSerTyrGluAlaGluIleValArgLeuThrThrProGlyPhe	562
Qy	419	TCACATCTCTGTCGATCAGTCAGCTGTCCTTTTATAAGTAAGTATAGTAACACAG	478
Db	563	SerHisSerCysSerMetSerGlnAsnPheAspMetPheValSerHisTyrSerSerVal	582
Qy	479	AAGAATCCACACTGTGTCTCTTTTGAAGTACTACTCGATTTTACATTTGATGGATGCTCTAC	538
Db	583	SerThrProProCysValHisValTyrLysLeuSerGlyProAspAspPheProLeuHis	602
Qy	539	AAACAAGGAATTTTGGCCACCATTTTGGATTTCAGCAGGTCTCTTCTCTGACTATACT	598
Db	603	LysGlnProArgPheTrpAlaSerMetMetGluAlaAlaSerCysProProAspTyrVal	622
Qy	599	CTCCAGAAATTTCTCTTTTGAAGTACTACTCGATTTTACATTTGATGGATGCTCTAC	658
Db	623	ProProGluIlePheHisPheHisThrArgSerAspValArgLeuTyrGlyMetIleTyr	642
Qy	659	AAGCCTCATGATCTACAGCTCGAAGAAATATCTACTGTGCTGTTTCATATATGTTGGT	718
Db	643	LysProHisAlaLeuGlnProGlyLysLysHisProThrValLeuPheValTyrGlyGly	662
Qy	719	CCTCAGCTCAGTTGGTGAATATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACC	778
Db	663	ProGlnValGlnLeuValAsnAsnSerPheLysGlyIleLysTyrLeuArgLeuAsnThr	682
Qy	779	CTAGCCTCTCTAGTTTATGTGTTAGTATAGCAACAGGGATCTCTGCACCGAGGG	838
Db	683	LeuAlaSerLeuGlyTyrAlaValValIleAspGlyArgGlySerCysGlnArgGly	702
Qy	839	CTTAATTTTCAAGCGCTTTTAAATATAAATGGGTCAAATAGAAATTTGACGATCAGGTG	898
Db	703	LeuArgPheGluGlyAlaLeuLysAsnGlnMetGlyGlnValGluIleGluAspGlnVal	722
Qy	899	GAAGACTCCAATATCTAGCTTCTCGATATGATTTTCATTGACTTGTGTTGGGCATC	958
Db	723	GluGlyLeuGlnPheValAlaGluLysTyrGlyPheIleAspLeuSerArgValAlaIle	742
Qy	959	CACGGCTGGTCTTATGAGGATACCTCTCCTGATGGCATTAATGCAGAGGTGCAGATATC	1018
Db	743	HisGlyTrpSerTyrGlyPheLeuSerLeuMetGlyLeuIleHisLysProGlnVal	762

```

QY 1019 TTCAGGTTCTCTATGCTGGGCCCCAGTCACCTGTGATCTTCTATGATACAGGATAC 1078
Db 763 PheLysValAlaIleAlaGlyAlaProValThrValTrpMetAlaTyrAspThrGlyTyr 782

QY 1079 ACG 1081
Db 783 Thr 783

RESULT 10
; Sequence 35, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junten, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 35
; LENGTH: 879
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-976-674-35

Alignment Scores:
Pred. No.: 3,61e-133 Length: 879
Score: 1201.00 Matches: 234
Percent Similarity: 55.1% Conservative: 53
Best Local Similarity: 44.9% Mismatches: 72
Query Match: 62.0% Indels: 162
Ds: 2 Gaps: 3

US-10-825-632-8 (1-1083) x US-09-976-674-35 (1-879)

QY 5 GAAGATCCAGATCAGCTGAGTCGCTACTTGTCTCCAAAGAAATTTGATAGATAT 64
Db 263 AspAspProLysSerAlaGlyValAlaThrPheValIleGlnGluPheAspArgPhe 282

QY 65 TCTGGCTATTGGTGGTGTCCAAAGCTGAAACAACTCCAGTGGTGGT---AAAATCTT 121
Db 283 ThrGlyTyrTrpTrpCysProThrAlaSerTrpGluGlySerGluGlyLeuLysThrLeu 302

QY 122 AGAATTCATATCAAGAAATGATGATCTGAGGTGGAATATTCTGTTACATCCCT 181
Db 303 ArgIleLeuTyrGluGluValAspGluSerGluValGluValIleHisValProSerPro 322

QY 182 ATGTTGGAACAAGAGGCGCAGATTCATTCCGTTATCCTAAACAGGATCAGCAATCCT 241
Db 323 AlaLeuGluGluArgLysThrAspSerTyrArgTyrProArgThrGlySerLysAsnPro 342

QY 242 AAAGTCACCTTTAAAGATGTCAGAAATATGATTGATGCTGAAGGAGGATCATAGTTGAT 301
Db 343 LysIleAlaLeuLysLeuAlaGluPheGlnThrAspSerGlnGlyLysIleValSerThr 362

QY 302 GAAGTCAGAGGCTG----- 316
Db 363 GlnGluLysGluLeuValGlnProPheSerSerLeuPheProLysValGluTyrIleAla 382

QY 316 ----- 316
Db 383 ArgAlaGlyTyrTrpThrArgAspGlyLysTyrAlaTrpAlaMetPheLeuAspArgProGln 402

QY 316 ----- 316
Db 403 GlnTrpLeuGlnLeuValLeuLeuProProAlaLeuPheIleProSerThrGluAsnGlu 422

```

```

QY 317 -----GTATAT 322
Db 423 GluGlnArgLeuAlaSerAlaArgAlaValProArgAsnValGlnProTyrValValTyr 442

QY 323 TTTGAA----- 328
Db 443 GluGluValThrAsnValTrpIleAsnValHisAspIlePheTyrProPheProGlnSer 462

QY 328 ----- 328
Db 463 GluGlyGluAspGluLeuCysPheLeuArgAlaAsnGluCysLysThrGlyPheCysHis 482

QY 328 ----- 328
Db 483 LeuTyrLysValThrAlaValLeuLysSerGlnGlyTyrAspTrpSerGluProPheSer 502

QY 328 ----- 328
Db 503 ProGlyGluAspGluPheLysCysProLysGluGluIleAlaLeuThrSerGlyGlu 522

QY 329 -----GGCACAAAGACTCCCTTTAGAGCATCAC 358
Db 523 TrpGluValLeuAlaArgHisGlySerLysGlyThrLysAspThrProLeuGluHisHis 542

QY 359 CTGTAGCTAGTCAGTTACCTAAATCCCTGGAGAGGTGACAAGCTGACCTGACCGTGGCTAC 418
Db 543 LeuTyrValValSerTyrGluAlaGlyGluIleValArgLeuThrThrProGlyPhe 562

QY 419 TCACATTTCTGTCATCAGTCAGCAGCTGTGACTCTTTTATAAGTAAAGTATAGTAAACAG 478
Db 563 SerHisSerCysSerMetSerGlnAsnPheAspMetPheValSerHisTyrSerSerVal 582

QY 479 AAGAATCCACTGTGTGCTCTTACAAGCTATCAAGTCTCGAAGATGACCCCACTTGC 538
Db 583 SerThrProProCysValHisValTyrLysLeuSerGlyProAspAspProLeuHis 602

QY 539 AAACAAAGGAATTTTGGGCCACCATTTGGATTTCAGCAGGTCTCTCTCCTCAGTATACT 598
Db 603 LysGlnProArgPheTrpAlaSerMetGluAlaAlaSerCysProProAspTyrVal 622

QY 599 CCTCAGAAATTTCTCTTTGAAAGTACTACTGGGATTTACATTTGATGGGATGCTCTAC 658
Db 623 ProProGluIlePheHisPheHisThrArgSerAspValArgLeuTyrGlyMetIleTyr 642

QY 659 AAGCCTCATGATCTACAGCCTGGAAAGAAATATCTCTACTGCTGCTGTTATATATCGTGT 718
Db 643 LysProHisAlaLeuGlnProGlyLysLysHisProThrValLeuPheValTyrGlyGly 662

QY 719 CCTCAGGTGAGTGGTGAATATCGTTTAAAGAGTCAAGTATTTCCGCTTGAATACC 778
Db 663 ProGlnValGlnLeuValAsnAsnSerPheLysGlyIleLysTyrLeuArgLeuAsnThr 682

QY 779 CTAGCCTCTCTAGTTTATGTTGTTGATGATAGACACACGGGGATCTGTCCAGGGG 838
Db 683 LeuAlaSerLeuGlyTyrAlaValValIleAspGlyArgGlySerCysGlnArgGly 702

QY 839 CTTAAATTTGAAGCCCTTTTAAATATAAATAGGTCAATAGAAATGACATCAGGTG 898
Db 703 LeuArgPheGluGlyAlaLeuLysAsnGlnMetGlyGlnValGluIleGluAspGlnVal 722

QY 899 GAAGGACTCCAATATCTAGCTTCTCGATATGATTTGATTTAGCTAGCTGTTGGGATC 958
Db 723 GluGlyLeuGlnPheValAlaGluLysTyrGlyPheIleAspLeuSerArgValAlaIle 742

QY 959 CAGGCTGTCTCTAGGAGTACCTCTCCCTGATGGCATTAATGACAGATTCAGATATC 1018
Db 743 HisGlyTyrSerTyrGlyPheLeuSerLeuMetGlyLeuIleHisLysProGlnVal 762

QY 1019 TTCAGGTTGCTATTGCTGGGCCCCAGTCACTCTGTGGATCTTCTTATGATACAGGATAC 1078
Db 763 PheLysValAlaIleAlaGlyAlaProValThrValTrpMetAlaTyrAspThrGlyTyr 782

```

QY 1079 ACG 1081 295 ----- 295
Db 783 Thr 783
RESULT 11
US-09-976-674-19
; Sequence 19, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 19
; LENGTH: 658
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-19
Alignment Scores:
Pred. No.: 4,73e-133 Length: 658
Score: 1199.50 Matches: 239
Percent Similarity: 58.0% Conservative: 0
Best Local Similarity: 58.0% Mismatches: 0
Query Match: 61.9% Indels: 173
DB: 2 Gaps: 1
US-10-825-632-8 (1-1083) x US-09-976-674-19 (1-658)
QY 2 GAAGAAGATCCAGATCAGCTGAGTGGCTACCTTTCTCTCCAAAGAAGATTGTATAGA 61
Db 244 GluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluPheAspArg 263
QY 62 TATCTGGCTATTGGTGGTGTCCAAAGCTGAAACAACTCCCGAGTGGTGTAAATCTTT 121
Db 264 TyrSerGlyTyrTipCysProLysAlaGluThrThrProSerGlyGlyLysIleLeu 283
QY 122 AGAATTCATATGACAAATGATCAATCTGAGTGGAAATTTATTCATGTTACATCCCT 181
Db 284 ArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisValThrSerPro 303
QY 182 ATGTTGAAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 241
Db 304 MetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThrAlaAsnPro 323
QY 242 AAAGTCACCTTTTAAGATGTCAGAAATAATGATGATGCTGAAGAGGATCATA----- 295
Db 324 LysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIleIleAspVal 343
QY 295 ----- 295
Db 344 IleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGluTyrIleAla 363
QY 295 ----- 295
Db 364 ArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAspArgSerGln 383
QY 295 ----- 295
Db 384 ThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGluAspVal 403
QY 295 ----- 295
Db 404 MetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeuIleIleTyr 423

QY 295 ----- 295
Db 424 GluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPheProGlnSer 443
QY 295 ----- 295
Db 444 HisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPheArgHisLeu 463
QY 295 ----- 295
Db 464 TyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGlyLeuPro 483
QY 295 ----- 295
Db 484 AlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSerGlyGluTrp 503
QY 296 ----- 322
Db 504 GluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgLeuValTyr 523
QY 323 TTTGAAGGCACCAAGAGACTCCCTTTTAGACATCACCTGTACGTACTGCTAGTTACGTAAT 382
Db 524 PheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSerTyrValAsn 543
QY 383 CCTGAGAGGTGACAAAGCTGACCTGACCTGCTTACTCACATTTCTTGCTGCATCAGT 442
Db 544 ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysLysSerGln 563
QY 443 CACTGTGACTCTTTTATAGTAAGTATAGTAACAGAGAATCCACATGTGTGTCTCTT 502
Db 564 HisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeu 583
QY 503 TACAAGCTATCAAGTCTCGAAGATGACCAACTTGGCAAAACAAAGGAATTTTGGCCACC 562
Db 584 TyrLysLeuSerSerProGluAspProThrCysLysThrLysGluPheTrpAlaThr 603
QY 563 ARTTTGGATTACAGAGTCTCTCTCTGACTATATCTCTCCAGAAATTTTCTCTTTTGA 622
Db 604 IleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePheSerPheGlu 623
QY 623 AGTACTACTGGATTTTACATTTGATGGATGCTCTACAGCCTCATGATCTACACCTGGA 682
Db 624 SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly 643
QY 683 AAGAAATATCTACTGCTGCTGCTCATATATGCTGT 718
Db 644 LysLysTyrProThrValLeuPheIleTyrGlyGly 655
RESULT 12
US-09-976-674-29
; Sequence 29, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 29
; LENGTH: 832
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-29
Alignment Scores:
Pred. No.: 4,74e-125 Length: 832

Db	583	GLnAspPheAspMetPheValSerHisTyrSerSerValSerThrProCysValHis	602
Qy	500	CTTTACAAGCTATCAAGTCCTGAAGATGACCAACTTGC AAAACAAAGGAATTTTGGGCC	559
Db	603	ValTyrLysLeuSerGlyProAspAspPheLeuHisLysGlnProA-gPheTTPala	622
Qy	560	ACATTTTGGATTCACGAGGTCCTCTTCTGACTATATCTCTCCAGAAAATTTTCTCTTTT	619
Db	623	SerMetMetGluAlaAlaSerCysProProAspTyrValProGluIlePheHisPhe	642
Qy	620	GAAAGTACTACTGGATTTCATATTGTATGGGATGCTTACAAGCTCATGATCTACAGCCT	679
Db	643	HisThrArgSerAspValArgLeuTyrGlyMetIleTyrLysProHisAlaLeuGlnPro	662
Qy	680	GGAAAGAAATATCTTACTTGCTCTCTCATATATGTGTGCTCTCAGGTGCAGTTGGTGAAT	739
Db	663	GlyLysLysHisProThrValLeuPheValTyrGlyGlyProGlnValGlnLeuValAsn	682
Qy	740	AATCCGTTTAAAGAGATCAAGTATTTCGCTTGAATACCTAGCCCTCTCTAGGTATGTG	799
Db	683	AsnSerPheLysGlyIleLysTyrLeuArgLeuAsnThrLeuAlaSerLeuGlyTyrAla	702
Qy	800	GTGTAGTATGATAGACACAGAGGATCTGTACCGAGGCTTAAATTTCAAGCGCCTTT	859
Db	703	ValValValIleAspGlyArgGlySerCysGlnArgGlyLeuArgPheGluGlyAlaLeu	722
Qy	860	AAATATAAATGGGTCAAAATAGAAATTCAGCATAGCTGAGGAGCTCCAATCTAGCT	919
Db	723	LysAsnGlnMetGlyGlnValGluIleGluAspGlnValGluGlyLeuGlnPheValAla	742
Qy	920	TCTCGATATGATTCATGTAGCTTAGATCTGTGGGATCCACGGCTGGTCTTATGAGGA	979
Db	743	GluLysTyrGlyPheIleAspLeuSerArgValAlaIleHisGlyTyrSerTyrGlyGly	762
Qy	980	TACTCTCCCTGATGGCATTAATGCAAGAGTCAGATPCTTCAGGTTGCTATTGCTGGG	1039
Db	763	PheLeuSerLeuMetGlyLeuIleHisLysProGlnValPheLysAlaGlnProLeuAla	782
Qy	1040	CCCCAGTCACTCTG 1054	
Db	783	TyrProProArgLeu 787	
RESULT 13			
US-09-976-674-31			
; Sequence 31, Application US/09976674			
; Patent No. 6844180			
; GENERAL INFORMATION:			
; APPLICANT: Qi, Steve			
; APPLICANT: Akinsanya, Karen			
; APPLICANT: Riviere, Pierre			
; APPLICANT: Junier, Jean-Louis			
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV			
; FILE REFERENCE: 70669			
; CURRENT APPLICATION NUMBER: US/09/976,674			
; CURRENT FILING DATE: 2001-10-12			
; PRIOR APPLICATION NUMBER: US 60/240,117			
; PRIOR FILING DATE: 2000-10-12			
; NUMBER OF SEQ ID NOS: 61			
; SOFTWARE: PatentIn version 3.1			
; SEQ ID NO 31			
; LENGTH: 832			
; TYPE: PRT			
; ORGANISM: Homo sapiens			
US-09-976-674-31			
Alignment Scores:			
Pred. No.: 4,74e-125			
Score: 1132.50			
Percent Similarity: 52.2%			
Best Local Similarity: 42.3%			
Query Match: 76			
Indels: 175			
Length: 832			
Macches: 222			
Conservative: 52			
Mismatches: 76			
Indels: 175			

[illegible]

Db 263 AspAspProLysSerAlaGlyValAlaThrPheValIleGlnGluPheAspArgPhe 282
QY 65 TCTGGCTATTGGTGGTCTCCAAAGCTGAAACAATCCAGTGGTGGT---AAAATTCTTT 121
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 121
283 ThrGlyTyrTrpCysProThrAlaSerTrpGluGlySerGluGlyLeuLysThrLeu 302
QY 122 AGAATCTATATCAAGAAATGATGAATCTGAGTGGAAATATTTCATGTTTACATCCCT 181
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 181
303 ArgIleuLysGluGluValAspGluSerGluValGluValIleHisValProSerPro 322
QY 182 ATGTTGAAACAAGGAGGCGAGATTCTTCGCTTATCTTAAACAGGTACAGCAATCTCT 241
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 241
323 AlaLeuGluGluLysThrAspSerTyrArgTyrProArgThrGlySerLysAsnPro 342
QY 242 AAAGTCACCTTTAAGATGTCAGAAATAATGATTCATGCTGAAGGAAGATCATAGTTGAT 301
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 301
343 LysIleAlaLeuLysLeuAlaGluPheGlnThrAspSerGlnGlyLysIleValSerThr 362
QY 302 GAAGTCAGAGGCGCTG----- 316
Db :|||:|:|:| 316
363 GlnGluLysGluLeuValGlnProPheSerSerLeuPheProLysValGluTyrIleAla 382
QY 316 ----- 316
383 ArgAlaGlyTrpThrArgAspGlyLysTyrAlaTrpAlaMetPheLeuAspArgProGln 402
QY 316 ----- 316
403 GlnTrpLeuGlnLeuValLeuLeuProProAlaLeuPheIleProSerThrGluAsnGlu 422
QY 317 -----GTATAT 322
423 GluGlnArgLeuAlaSerAlaArgAlaValProArgAsnValGlnProTyrValValTyr 442
QY 323 TTTGAA----- 328
443 GluGluValThrAsnValTrpIleAsnValHisAspIlePheTyrProPheProGlnSer 462
QY 328 ----- 328
463 GluGlyGluAspGluLeuCysPheLeuArgAlaAsnGluCysLysThrGlyPheCysHis 482
QY 328 ----- 328
483 LeuTyrLysValThrAlaValLeuLysSerGlnGlyTyrAspTrpSerGluProPheSer 502
QY 328 ----- 328
503 ProGlyGluAspGluPheLysCysProIleLysGluGluIleAlaLeuThrSerGlyGlu 522
QY 329 -----GCCACCAAGACTCCCTTTAGAGCATCAC 358
523 TrpGluValLeuAlaArgHisGlySerLysGlyThrLysAspThrProLeuGluHisHis 542
QY 359 CTGTACGTAGTCAGTTAGCTAAATCCTGGAGAGGTGCAAGGCTGACTGACCGTGGCTAC 418
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 418
543 LeuTyrValValSerTyrGluAlaAlaGlyLysIleValArgLeuThrThrProGlyPhe 562
QY 419 TCACATCTTGTCGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTCAG 478
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 478
563 SerHisSerCysSerMetSerGlnAsnPheAspMetPheValSerHisTyrSerSerVal 582
QY 479 AAGAAATCCACACTGTGTCTCCCTTTACAAAGTATCAAGTCTCAAGATCAACCAACTGTC 538
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 538
583 SerThrProProCysValHisValTyrLysLeuSerGlyProAspAspProLeuHis 602
QY 539 AAAACAAAGGAATTTGGGCGACCATTTTGGATTTCAGCAGGCTCTCTTCGACTACTACT 598
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 598
603 LysGlnProArgPheTrpAlaSerMetMetGluAlaAlaSerCysProProAspTyrVal 622
QY 599 CCTCAGAAATTTCTCTTTTGAAGTACTACTGAGTTTACATTTGATGGGATGCTCTAC 658
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 658
623 ProProGluIlePheHisPheHisThrArgSerAspValArgLeuTyrGlyMetIleTyr 642

QY 659 AAGCTCATGATCTACAGCTGGAAGAAATATCTCTACTGTCTGTCTCATATATGGTGGT 718
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 718
643 LysProHisAlaLeuGlnProGlyLysHisProThrValLeuPheValTyrGlyGly 662
QY 719 CTTAGGTGAGTGGTGGTGAATATCGTTTAAAGGAGTCAAGTATTTCCGCTTCAATACC 778
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 778
663 ProGlnValGlnLeuValAsnAsnSerPheLysGlyIleLysTyrLeuArgLysThr 682
QY 779 CTAGCTCTCTAGTTTATGTGTTTGTAGTGATAGACACAGGGATCTGTCCACCGAGGG 838
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 838
683 LeuAlaSerLeuGlyTyrAlaValValIleAspGlyArgGlySerCysGlnArgGly 702
QY 839 CTTAAATTTGAAGGCGCTTTTAAATATAAATGGTCAATAGAAATGACGATCAGGTG 898
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 898
703 LeuArgPheGluGlyAlaLeuLysAsnGlnMetGlyGlnValGluIleGluAspGlnVal 722
QY 899 GAAGGACTCCAATATCTAGCTTCTCGATATGATTTTCAATGATCTGTGGGGCATC 958
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 958
723 GluGlyLeuGlnPheValAlaGluLysTyrGlyPheIleAspLeuSerArgValAlaIle 742
QY 959 CACGGCTGGTCTATGGAGGATACCTCTCCCTGATGCGATTAAATGACAGGTCAGATATC 1018
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 1018
743 HisGlyTyrSerTyrGlyPheLeuSerLeuMetGlyLeuIleHisLysProGlnVal 762
QY 1019 TTTACGGCTTGCTATTGTGGGCGCCAGTCACCTCTG 1054
Db :|||||:| 1054
763 PheLysAlaGlnProLeuAlaTyrProProArgLeu 774

RESULT 15
US-09-976-674-39
; Sequence 39, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Oi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 706693
; CURRENT APPLICATION NUMBER: US/09/976,674
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 39
; LENGTH: 819
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-39

Alignment Scores:
Pred. No.: 9,69e-123 Length: 819
Score: 1113.00 Matches: 220
Percent Similarity: 52.9% Conservative: 51
Best Local Similarity: 43.0% Mismatches: 79
Query Match: 57.4% Indels: 162
DB: 2 Gaps: 3

US-10-825-632-8 (1-1083) x US-09-976-674-39 (1-819)

QY 5 GAAGATGCCAGATCAGCTGGAGTCCGTACTTCTCTCCAAAGAAATTTGATAGATAT 64
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 64
263 AspAspProLysSerAlaGlyValAlaThrPheValIleGlnGluPheAspArgPhe 282
QY 65 TCTGGCTATTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 121
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 121
283 ThrGlyTyrTrpTyrCysProThrAlaSerTrpGluGlySerGluGlyLeuLysThrLeu 302
QY 122 AGAATCTATATGAAGAAATGATGATCTGAGTGGAAATATTTCATGTTTACATCCCT 181
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 181
303 ArgIleuTyrGluGluValAspGluSerGluValGluValIleHisValProSerPro 322

QY	182	ATGTTGAAACAAAGGAGGCGAGATTTCATTCGTTATCTCTAAACAGGTACAGCAATCCT	241	QY	779	CTAGCCTCTCTAGGTTATGTGTTAGTATAGACAAACAGGCGATCCTGTCCAGGAGG	838
Db	323	AlaLeuGluGluArgLysThrAspSerTyrArgTyrProArgThrGlySerLysAsnPro	342	Db	683	LeuAlaSerLeuGlyTyrAlaValValIleAspGlyArgGlySerCysGlnArgGly	702
QY	242	AAAGTCACCTTTTACAGATCTCAAAATATGATGATCTCAAGAGGAGGATCATAGTTGAT	301	QY	839	CTTAAATTTGAAAGGCGCCTTTAAATATATAAATGGGTCAAATAGAAATTTGACGATCAGGTG	898
Db	343	LysIleAlaLeuLysLeuAlaGluPheGlnThrAspSerGlnGlyLysIleValSerThr	362	Db	703	LeuArgPheGluGlyAlaLeuLysAsnGlnMetGlyGlnValGluIleGluAspGlnVal	722
QY	302	GAAGTCAGAGAGGCTG-----	316	QY	899	GAAGGACTCAATATCTAGCTTCTCGATATGATTTTCATTGACTTGTAGATCGGTGGGCATC	958
Db	363	GlnGluLysGluLeuValGlnProPheSerSerLeuPheProLysValGluTyrIleAla	382	Db	723	GluGlyLeuGlnPheValAlaGluLysTyrGlyPheIleAspLeuSerArgValAlaIle	742
QY	316	-----	316	QY	959	CACGGCTGTCTTATGGAGGATACCTCTCCCTGTAGTGGCATTAAATCGAGAGGTGAGATATC	1018
Db	383	ArgAlaGlyTyrThrArgAspGlyLysTyrAlaTrpAlaMetPheLeuAspArgProGln	402	Db	743	HisGlyTrpSerTyrGlyGlyPheLeuSerLeuMetGlyLeuIleHisLysProGlnVal	762
QY	316	-----	316	QY	1019	TTCAGGGTTGCTATTGCTGGGCGCCCACTGCTCTG	1054
Db	403	GlnTrpLeuGlnLeuValLeuProProAlaLeuPheIleProSerThrGluAsnGlu	422	Db	763	PheLysAlaGlnProLeuAlaTyrProProArgLeu	774
QY	317	-----	317				
Db	423	GluGlnArgLeuAlaSerAlaArgAlaValProArgAsnValGlnProTyrValValTyr	442				
QY	323	TTTGRA-----	328				
Db	443	GluGluValThrAsnValTrpIleAsnValHisAspIlePheTyrProPheProGlnSer	462				
QY	328	-----	328				
Db	463	GluGlyGluAspGluLeuCysPheLeuArgAlaAsnGluCysLysThrGlyPheCysHis	482				
QY	328	-----	328				
Db	483	LeuTyrLysValThrAlaValLeuLysSerGlnGlyTyrAspTrpSerGluProPheSer	502				
QY	328	-----	328				
Db	503	ProGlyGluAspGluPheLysCysProIleLysGluGluIleAlaLeuThrSerGlyGlu	522				
QY	329	-----	329				
Db	523	TrpGluValLeuAlaArgHisGlySerLysGlyThrLysAspThrProLeuGluHisHis	542				
QY	359	CTGTACGTAGTCAGTTTACCTAAATCCTCGAGAGGTGACAAAGCTGACTGACCGTGGCTAC	418				
Db	543	LeuTyrValValSerTyrGluAlaAlaGlyGluIleValArgLeuThrThrProGlyPhe	562				
QY	419	TCACATCTCTGTCATCAGTCAGCAGCTGTGACTCTTTTATAGTAGTATAGTAACCAG	478				
Db	563	SerHisSerCysSerMetSerGlnAsnPheAspMetPheValSerHisTyrSerSerVal	582				
QY	479	AGAATCCACACTGTGTGCTCCTTTACAAGCTATCAAGTCTCTGAAGATGACCCAACTTGC	538				
Db	583	SerThrProCysValHisValTyrLysLeuSerGlyProAspAspProLeuHis	602				
QY	539	AAACAAAGAAATTTGGGCCACCAATTTGGATTACAGAGGTCTCTTCTCTGACTATACT	598				
Db	603	LysGlnProArgPheTrpAlaSerMetMetGluAlaAlaSerCysProProAspTyrVal	622				
QY	599	CTCCAGAAATTTCTCTTCAAGTACTACTGTGATTTACATTTGATGGATGCTCTAC	658				
Db	623	ProGluIlePhePheHisPheHisThrArgSerAspValArgLeuTyrGlyMetIleTyr	642				
QY	659	AAGCCTCATGATCTACAGCTCGAAAGAAATATCTACTGTGCTGTTTCATATATGGTGGT	718				
Db	643	LysProHisAlaLeuGlnProGlyLysLysHisProThrValLeuPheValTyrGlyGly	662				
QY	719	CCTCAGGTGAGTGGTGAATATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACC	778				
Db	663	ProGlnValGlnLeuValAsnAsnSerPheLysGlyIleLysTyrLeuArgLeuAsnThr	682				

Search completed: May 2, 2006, 01:42:33
Job time : 58.1838 secs

THIS PAGE BLANK (USPTO)

Sequence 2, Appli
Sequence 7, Appli
Sequence 7, Appli
Sequence 620, App
Sequence 7, Appli
Sequence 3, Appli
Sequence 619, App
Sequence 3, Appli
Sequence 23, Appli
Sequence 27, Appli
Sequence 23, Appli
Sequence 27, Appli
Sequence 12, Appli
Sequence 22, App
Sequence 226, App
Sequence 11, Appli
Sequence 11, Appli
Sequence 4, Appli
Sequence 33, Appli
Sequence 35, Appli
Sequence 33, Appli
Sequence 35, Appli
Sequence 19, Appli
Sequence 19, Appli
Sequence 29, Appli
Sequence 31, Appli
Sequence 29, Appli
Sequence 31, Appli
Sequence 37, Appli
Sequence 39, Appli
Sequence 37, Appli
Sequence 39, Appli
Sequence 4, Appli

12	1495	77.1	580	6	US-11-140-224-2
13	1226	63.3	690	3	US-09-976-674-7
14	1226	63.3	690	5	US-10-982-512-7
15	1220.5	63.0	889	4	US-10-072-012-620
16	1220.5	63.0	830	4	US-10-415-122-7
17	1220.5	63.0	863	3	US-09-976-674-3
18	1220.5	63.0	863	4	US-10-072-012-619
19	1220.5	63.0	863	5	US-10-982-512-3
20	1220.5	63.0	892	3	US-09-976-674-23
21	1220.5	63.0	892	3	US-09-976-674-27
22	1220.5	63.0	892	5	US-10-982-512-23
23	1220.5	63.0	892	5	US-10-982-512-27
24	1220.5	63.0	892	5	US-10-433-757-12
25	1212.5	62.6	863	4	US-10-072-012-224
26	1212.5	62.6	863	4	US-10-072-012-226
27	1212.5	62.6	863	3	US-09-976-674-11
28	1209	62.4	661	5	US-10-982-512-11
29	1209	62.4	661	5	US-10-415-122-4
30	1204.5	62.2	869	4	US-09-976-674-33
31	1201	62.0	879	3	US-09-976-674-35
32	1201	62.0	879	5	US-10-982-512-33
33	1201	62.0	879	5	US-10-982-512-35
34	1201	62.0	879	5	US-09-976-674-19
35	1199.5	61.9	658	5	US-10-982-512-19
36	1199.5	61.9	658	5	US-09-976-674-29
37	1132.5	58.4	832	3	US-09-976-674-31
38	1132.5	58.4	832	5	US-10-982-512-29
39	1132.5	58.4	832	5	US-10-982-512-31
40	1132.5	58.4	832	5	US-09-976-674-37
41	1113	57.4	819	3	US-09-976-674-39
42	1113	57.4	819	5	US-10-982-512-37
43	1113	57.4	819	5	US-10-982-512-39
44	1113	57.4	819	5	US-10-982-512-39
45	1086.5	56.1	310	3	US-09-993-959-4

ALIGNMENTS

RESULT 1
US-10-825-632-7
; Sequence 7, Application US/10825632
; Publication No. US20040191826A1
; GENERAL INFORMATION:
; APPLICANT: ABBOTT, Catherine Anne
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES
; FILE REFERENCE: FCSB-100-Div. 1
; CURRENT APPLICATION NUMBER: US/10/825,632
; CURRENT FILING DATE: 2004-04-15
; PRIOR APPLICATION NUMBER: US 10/070,464
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: PCT/AU00/01085
; PRIOR FILING DATE: 2000-09-11
; PRIOR APPLICATION NUMBER: AU PQ5709
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: AU PQ2762
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 360
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-825-632-7

Alignment Scores:			
Pred. No.:	6,53e-192	Length:	360
Score:	1933.00	Matches:	360
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	99.7%	Indels:	0
DB:	4	Gaps:	0

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: May 2, 2006, 02:43:09 ; Search time 31.0173 Seconds
(without alignments)
2917.786 Million cell updates/sec

Title:
Perfect score:
Sequence:
1 ggaagaagatgcagatcag.....tatgatacaggatacacgga 1083

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 1867569 seqs, 417829326 residues
Total number of hits satisfying chosen parameters: 3735138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame_n2p_model -DEV=xlp
-Q=/abs/ABSWEB_spool/US10825632/runat_01052006_105957_3449/app_query.fasta_1
-DB=Published Applications AA Main -OPMT=fastan -SUFFIX=n2p.rapbm
-MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1
-MATRIX=blosum62 -TRANS=human40.cdi -LIST=45 -DOALIGN=200 -THR SCORE=pct
-THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext
-HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs602p
-USER=US10825632 @CGN 1 1 503 @runat 01052006 105957_3449 -NCPU=6 -ICPU=3
-NO MAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
1: /cn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
3: /cn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
4: /cn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
5: /cn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
6: /cn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1933	99.7	360	4	US-10-825-632-7
2	1836.5	94.8	882	3	US-09-976-674-1
3	1836.5	94.8	882	4	US-10-054-776-2
4	1836.5	94.8	882	4	US-10-170-789-38
5	1836.5	94.8	882	4	US-10-311-035-9
6	1836.5	94.8	882	4	US-10-072-012-622
7	1836.5	94.8	882	4	US-10-415-122-6
8	1836.5	94.8	882	4	US-10-825-632-1
9	1836.5	94.8	882	5	US-10-982-512-1
10	1756.5	90.6	883	4	US-10-072-012-621
11	1495	77.1	580	4	US-10-275-505-2

US-10-825-632-8 (1-1083) x US-10-825-632-7 (1-360)

```
QY 2 GAAGAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAGAGATTTGATAGA 61
Db 1 GluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluPheAspArg 20
QY 62 TATTCTGGCTATTGGTGGTGTCCAAAGCTGAAACCACTCCAGTGGTGTGTAATCTTT 121
Db 21 Ty-SerGlyTyrTrpTrpCysProLysAlaGluThrProSerGlyGlyLysLeu 40
QY 122 AGAATCTTATATGAAGAAATCATGATCTGAGTGGAAATTTATTCATGTTACATCCCT 181
Db 41 ArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleHisValThrSerPro 60
QY 182 ATGTTGGAACAGGAGGAGGAGATTCATTCGTTATCTTAAACAGGTACAGAAATCTT 241
Db 61 MetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThrAlaAsnPro 80
QY 242 AAAGTCACATTTTAAGATGTCAGAAATATGATGCTGAAGAAAGGATCATAGTTGAT 301
Db 81 LysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIleValAsp 100
QY 302 GAAGTCAGAGGCTGGTATATTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTG 361
Db 101 GluValArgGluValTyrPheGluGlyThrLysAspSerProLeuGluHisLeu 120
QY 362 TAGCTAGTCAGTTACGTAATCTCTGGAGAGTGACAGGCTGACTGACCGTGGCTACTCA 421
Db 121 TyrValValSerTyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSer 140
QY 422 CATCTCTGCTGATCAGTCAGCACTGTGACTTCTTTATAGTAAGTATAGTAACAGAG 481
Db 141 HisSerCysCysIleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLys 160
QY 482 AATCACACTGTGTGCTCTTTACAAGCTATCAAGTCTCAAGTCTCAAGTCACTTGCAGAA 541
Db 161 AsnProHisCysValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLys 180
QY 542 ACAAGGAAATTTGGGACCACTTTTGATTTCAGCAGGTCCTCTTCCTGACTATATCTCT 601
Db 181 ThrLysGluPheThrAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrPro 200
QY 602 CCAGAAATTTTCTCTTTGAAAGTACTACTGGAATTTACATTTGATGGGATGCTCTACAG 661
Db 201 ProGluIlePheSerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLys 220
QY 662 CCTCATGATCTACAGCCTGGAAGAAATATCTCTGCTGCTGCTTCATATATGTTGGTCTCT 721
Db 221 ProHisAspLeuGlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyPro 240
QY 722 CAGGTGCAGTTGGTGAATATCGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCCCTA 781
Db 241 GlnValGlnLeuValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeu 260
QY 782 GCCTCTCTAGTTATGTTGTTAGTAGATAGACACAGGGGATCCTGTCACCGAGGCTT 841
Db 261 AlaSerLeuGlyTyrValValValValIleAspAsnArgGlySerCysHisArgGlyLeu 280
QY 842 AAATTTGAGGCCCTTTAAATATAAATGGTCAAAATAGAAATTCAGATCAGGTGGAA 901
Db 281 LysPheGluGlyAlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGlu 300
QY 902 GGACTCCATATCTAGCTTCTCGATATGATTTTCATTTGACTTAGATCGTGGGATCCAC 961
Db 301 GlyLeuGlnTyrLeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHis 320
QY 962 GGCTGCTCTATGGAGGATPACTCTCCCTGATGGCATTATATGACAGGTCAGATATCTTC 1021
Db 321 GlyTrpSerTyrGlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePhe 340
QY 1022 AGGTTGCTATTGCTGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACAG 1081
Db 341 ArgValAlaIleAlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThr 360
```

RESULT 2
US-09-976-674-1
; Sequence 1, Application US/09976674
; Patent No. US20020115843A1
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; FILE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 882
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-1

Alignment Scores:

Pred. No.:	9,51e-182	Length:	882
Score:	1836.50	Matches:	360
Percent Similarity:	67.5%	Conservative:	0
Best Local Similarity:	67.5%	Mismatches:	0
Query Match:	94.8%	Indels:	173
DB:	3	Gaps:	1

US-10-825-632-8 (1-1083) x US-09-976-674-1 (1-882)

```
QY 2 GAAGAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAGAGATTTGATAGA 61
Db 244 GluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluPheAspArg 263
QY 62 TATTCTGGCTATTGGTGGTGTCCAAAGCTGAAACCACTCCAGTGGTGTGTAATCTTT 121
Db 264 TyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGlyLysLeu 283
QY 122 AGAATCTTATATGAAGAAATGATGAATCTGAGTGGAAATTTATTCATGTTACATCCCT 181
Db 284 ArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleHisValThrSerPro 303
QY 182 ATGTTGGAACAGGAGGAGGAGATTCATTCGTTATCTTAAACAGGTACAGAAATCTT 241
Db 304 MetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThrAlaAsnPro 323
QY 242 AAAGTCACATTTTAAGATGTCAGAAATATGATGCTGAAGAAAGGATCATGTA 295
Db 324 LysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIleValAspVal 343
QY 295 ----- 295
Db 344 IleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGluTyrIleAla 363
QY 295 ----- 295
Db 364 ArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAspArgSerGln 383
QY 295 ----- 295
Db 384 ThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGluAspVal 403
QY 295 ----- 295
Db 404 MetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeuIleIleTyr 423
QY 295 ----- 295
Db 424 GluGluThrAspIleTrpIleAsnIleHisAspIlePheHisValPheProGlnSer 443
```

us-10-825-632-8.n2p.rapbm

Mon May 8 13:41:59 2006

GENERAL INFORMATION:
; APPLICANT: Mark Robert Edbrooke
; APPLICANT: Alan Peter Lewis
; TITLE OF INVENTION: NOVEL PROTEIN
; FILE REFERENCE: QG1042US
; CURRENT APPLICATION NUMBER: US/10/054,776
; CURRENT FILING DATE: 2002-01-23
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 882
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-054-776-2

Alignment Scores: 9,51e-182 Length: 882
Pred. No.: 1836,50 Matches: 360
Score: 1836,50 Conservative: 0
Percent Similarity: 67,5% Mismatches: 0
Best Local Similarity: 67,5% Indels: 173
Query Match: 94,8% Gaps: 1
DB: 4

US-10-825-632-8 (1-1083) x US-10-054-776-2 (1-882)

QY 2 GAAGAGATCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAGAGAGAAATTTGATAGA 61
Db 244 GluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluPheAspArg 263
QY 62 TATTCTCGCTATTGGTGGTGTCCAAAGCTGAAACAACTCCAGTGGTGGTAAAATTCCT 121
Db 264 TyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGlyLysLeu 283
QY 122 AGAATTCCTATATGAAGAAATGATGAATCTGAGGTGGAATTTATTCATGTTACATCCCT 181
Db 284 ArgLeuTyrGluGluAsnAspGluSerGluValGluLeuLeuLeuLeuLeuLeuLeu 303
QY 182 ATGTGGGAAACAAGAGCGCAGATTCATTCCTGTTATCTTAAACAGGTACAGCAATCCT 241
Db 304 MetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThrAlaAsnPro 323
QY 242 AAGTCACTTTAAGATGTGAGAAATATGATGATGCTGAGTCTGAAGAGGATCATA----- 295
Db 324 LysValThrPheLysMetSerGluLeuMetIleAspAlaGluGlyArgIleIleAspVal 343
QY 295 ----- 295
Db 344 IleAspLysGluLeuLeuGlnProPheGluIleLeuPheGluGlyValGluTyrIleAla 363
QY 295 ----- 295
Db 364 ArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAspArgSerGln 383
QY 295 ----- 295
Db 384 ThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGluAspVal 403
QY 295 ----- 295
Db 404 MetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeuIleIleTyr 423
QY 295 ----- 295
Db 424 GluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPheProGlnSer 443
QY 295 ----- 295
Db 444 HisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPheArgHisLeu 463
QY 295 ----- 295
Db 464 TyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGlyGlyLeuPro 483

295 ----- 295
444 HisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPheArgHisLeu 463
295 ----- 295
464 TyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGlyGlyLeuPro 483
295 ----- 295
484 AlaProSerAspPheLysCysProIleLysGluIleAlaIleThrSerGlyGluTrp 503
296 -----GTTGATGAAGTCAGAAAGCTGGGTATAT 322
504 GluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgLeuValTyr 523
323 TTGGAAGGACCAAGAGACTCCCTTTAGAGCATCACCTGTACGTAGTACGTAAAT 382
524 PheGluGlyThrLysAspSerProLeuGluHisLysLeuTyrValValSerTyrValAsn 543
383 CTTGAGAGGTCAGAGGCTGACTGACCGTGGCTACTCACATTCCTTCTGCTCATCAGTCAG 442
544 ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysLysSerGln 563
443 CACTGTGACTCTTTATAGTAAGTATAGTAACCAAGAGAAATCCACACTGTGTGTCCTT 502
564 HisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeu 583
503 TACAAGCTATCAAGTCTGAGATGACCAACTTCGAAACAAAGGAATTTGGGCCAC 562
584 TyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPheTrpAlaThr 603
563 ATTTTGATTCACAGTCTCTCTGACTATATCTCTCCAGAAATTTCTTTTGAA 622
604 IleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePheSerPheGlu 623
623 AGTACTACTGATTTACATTTGATGGATGCTCTACAGCTCATGATCAGCTGGA 682
624 SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly 643
683 AAGAAATATCTACTGCTGCTTCATATATGTTGGTCTCAGTGTGAGTGTGATTAAT 742
644 LysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeuValAsnAsn 663
743 CGCTTTAAGAGAGTCAAGATTTCCGCTTGAATACCTAGCTCTCTAGTTTATGCTGT 802
664 ArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrValVal 683
803 GTAGTATAGACAACAGGGGATCCTGTCCCGGGCTTAAATTTGAAGCGCCTTTAAA 862
684 ValValIleAspAsnArgLysSerCysHisArgGlyLeuLysPheGluGlyAlaPheLys 703
863 TATTAATGGTCAATAGAATTAAGATGACGATCAGGTGGAAGACTCCAAATATCTAGCTTCT 922
704 TyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyrLeuAlaSer 723
923 CGATATGATTTCAATGACTTATAGTGTGGGCTCCACCGCTGGTCTTATGAGGATAC 982
724 ArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyrGlyGlyTyr 743
983 CTCTCCCTGATGCTATTAATGAGAGGTACAGATATCTTCAGGGTGTCTATTCGTGGGCC 1042
744 LeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAlaGlyAla 763
1043 CCAGTCACTCTGTGGATCTTCTATGATACAGATACAG 1081
764 ProValThrLeuTrpIlePheTyrAspThrGlyTyrThr 776

RESULT 3
US-10-054-776-2
; Sequence 2, Application US/10054776
; Publication No. US20030165818A1

```
QY 295 ----- 295
Db 484 AlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSerGlyGluTyr 503
QY 296 -----GTTGATGAGCTCAAGAGCTGGTATAT 322
Db 504 GluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgLeuValTyr 523
QY 323 TTTGAAGGACCAAGAGCTCCCTTTAGAGCATCACTGTACGTAGTCAGTTACGTAAAT 382
Db 524 PheGluGlyThrLysAspSerProLeuGluHisIleuTyValValSerTyValAsn 543
QY 383 CTGGAGAGGTCACAGGCTGACGCGCTGCTACTCACTTCTGCTGCATCAGTAC 442
Db 544 ProGlyGluValThrArgLeuThrAspArgGlyTySerHisSerCysValSerGln 563
QY 443 CACTGTGACTCTTTTATAGTAACTATAGTACAGAGAAATCCACACTGTGTGCTCT 502
Db 564 HisCysAspPhePheIleSerLysTySerAsnGlnLysAsnProHisCysValSerLeu 583
QY 503 TACAAGCTATCAAGTCTGAGATGACCAACTTGCAAAACAAGAAATTTTGGGCCACC 562
Db 584 TyrLysLeuSerSerProGluAspProThrCysLysThrLysGluPheTrpAlaThr 603
QY 563 ATTTTGGATTCCAGAGTCCCTTCTCTGACTATACCTCCAGAAATTTCTCTTTGAA 622
Db 604 IleLeuAspSerAlaGlyProLeuProAspTyThrProProGluIlePheSerPheGlu 623
QY 623 AGTACTACTGATTTACATTTGATGGATGCTCTACAAGCCTCATGATCTACAGCTGGA 682
Db 624 SerThrThrGlyPheThrLeuTyGlyMetLeuTyLysProHisAspLeuGlnProGly 643
QY 683 AAGAATATCTACTGCTGCTTATATATGATGGTCTCTCAAGCCTCATGATCTACAGCTGGA 742
Db 644 LysLysTyProThrValLeuPheIleTyGlyGlyProGlnValGlnLeuValAsnAsn 663
QY 743 CGGTTAAAGAGTCAAGTATTTCCGCTTGAATACCTAGCTCTCTAGGTTAGTGTT 802
Db 664 ArgPheLysGlyValLysTyPheArgLeuAsnThrLeuAlaSerLeuGlyTyValVal 683
QY 803 GTAGTATAGAACACAGGGGATCCGTGTCACCGAGGGCTTAAATTTGAAGCGCTTAA 862
Db 684 ValValIleAspAsnArgLysSerCysHisArgGlyLeuLysPheGluGlyAlaPheLys 703
QY 863 TATAAATGGTCAATAGAAATTCAGCATCAGGTGGAAGGACTCCCAATATCTAGTTCT 922
Db 704 TyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyLeuAlaSer 723
QY 923 CGATATGATTTCACTGACTTATAGATCGTGTGGGCATCCACGCTGCTATGAGAGATAC 982
Db 724 ArgTyAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyGlyGlyTy 743
QY 983 CTCTCCCTGATGGCATTATGACAGGTTCAGATATCTTCAGGTTCTCTATTCCTGGGCC 1042
Db 744 LeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAlaGlyAla 763
QY 1043 CCAGTCACTGTGTGATCTTCTATGATACAGGATACAG 1081
Db 764 ProValThrLeuTyrIlePheTyAspThrGlyTyThr 776

RESULT 4
US-10-170-789-38
; Sequence 38, Application US/10170789
; Publication No. US20030180930A1
; GENERAL INFORMATION:
; APPLICANT: Rachel E. Meyers
; APPLICANT: Olandt, Peter J.
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Curtis, Rory A. J.
; APPLICANT: Williamson, Mark
; APPLICANT: Welch, Nadine
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASE, PHOSPHATASE,
; AND PROTEASE FAMILY MEMBERS AND USES THEREOF

FILE REFERENCE: 10448-191001
; CURRENT APPLICATION NUMBER: US/10/170,789
; CURRENT FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: US 09/797,039
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: PCT/US01/06525
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/186,061
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 09/882,166
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: PCT/US01/19269
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/212,078
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 09/934,406
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: PCT/US01/26052
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/226,740
; PRIOR FILING DATE: 2000-08-21
; PRIOR APPLICATION NUMBER: US 09/861,801
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: PCT/US01/16549
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: US 60/205,508
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: US 09/801,267
; PRIOR FILING DATE: 2001-03-06
; PRIOR APPLICATION NUMBER: PCT/US01/07138
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: US 60/187,454
; PRIOR FILING DATE: 2000-03-07
; PRIOR APPLICATION NUMBER: US 09/829,671
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: PCT/US01/40483
; PRIOR FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: US 60/197,508
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: US 09/961,721
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: PCT/US01/29904
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: US 60/235,023
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US 10/045,367
; PRIOR FILING DATE: 2001-11-07
; PRIOR APPLICATION NUMBER: US 60/246,561
; PRIOR FILING DATE: 2000-11-07
; PRIOR APPLICATION NUMBER: US 09/801,275
; PRIOR FILING DATE: 2001-03-06
; PRIOR APPLICATION NUMBER: PCT/US01/07074
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: US 60/187,420
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 38
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-170-789-38

Alignment Scores:
Pred. No.: 9,51e-182 Length: 882
Score: 1836.50 Matches: 360
Percent Similarity: 67.5% Conservative: 0
Best Local Similarity: 67.5% Mismatches: 0
Query Match: 94.8% Indels: 173
DB: 4 Gaps: 1

US-10-825-632-8 (1-1083) x US-10-170-789-38 (1-882)
```


QY 2 GAAAGATGCCAGATCAGTCGAGTCGCTACCTTGTCTCCAAAGAAATTTGTAGTA 61
DB 244 GluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGluPheAspArg 263
QY 62 TATTCTGGCTATTGGTGGTCCAAAGCTGAAACAACTCCAGTGGTGTAAATCTT 121
DB 264 TyrSerGlyTyrTrpCysProLysAlaGluThrTrpSerGlyGlyLysIleLeu 283
QY 122 AGAATCTATATCAAGAAATGATGATCTGAGGTGGAAATTTATTCATGTTACATCCCT 181
DB 284 ArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleHisValThrSerPro 303
QY 182 ATGTGGAAACAAGGAGGCGAGATTCATTCCTGTTATCTCTAAACAGGTACAGCAATCT 241
DB 304 MetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThrAlaAsnPro 323
QY 242 AAAGTCACTTTTAAAGATGTCAGAAATAATGATGCTGAGGAGGAGCATCAT- 295
DB 324 LysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIleAspVal 343
QY 295 ----- 295
DB 344 IleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGluTyrIleAla 363
QY 295 ----- 295
DB 364 ArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAspArgSerGln 383
QY 295 ----- 295
DB 384 ThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGluAspAspVal 403
QY 295 ----- 295
DB 404 MetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeuIleIleTyr 423
QY 295 ----- 295
DB 424 GluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPheProGlnSer 443
QY 295 ----- 295
DB 444 HisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPheArgHisLeu 463
QY 295 ----- 295
DB 464 TyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGlyLeuPro 483
QY 295 ----- 295
DB 484 AlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSerGlyGluTrp 503
QY 296 -----GTTGATGAAGTCAGAAGCTGGTATAT 322
DB 504 GluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgGlyLeuValTyr 523
QY 323 TTGGAAGGACCAAAAGACTCCCTTTAGAGCATCACTGTAGTCAGTCAGTACGTTAAAT 382
DB 524 PheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSerTyrValAsn 543
QY 383 CTGGAGAGGTCAAGGCTGACTGACCGTGGCTACTCACATTTCTGTCAGTCAGTCAG 442
DB 544 ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysGlyIleSerGln 563
QY 443 CACTGTGACTCTTTTATAAGTATAGTATAACCAAGAAATCCACACTGTGTGTCCTT 502
DB 564 HisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeu 583
QY 503 TACAAGCTATCAGTCTCAAGATGACCACTTCGCAAAACAAGGAATTTTGGCCACC 562
DB 584 TyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPheTrpAlaThr 603
QY 563 ATTTTGGATTTCAGCAGGTCTCTCTCTGACTATATCTCTCCAGAAATTTTCTCTTTTGA 622

DB 604 IleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePheSerPheGlu 623
QY 623 AGTACTACTGATTTACATTTGATGGATGCTCTACAGCCTCATGATCTACAGCCTGA 682
DB 624 SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly 643
QY 683 AAGAAATATCTACTGTCTGTTTCATATATGTTGGTCCCTCAGGTGCAGTTGGTGAATAAT 742
DB 644 LysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeuValAsnAsn 663
QY 743 CGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCTAGCCTCTCTAGGTTATGGTGT 802
DB 664 ArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrValVal 683
QY 803 GTAGTGATAGACAACACGGGATCTGTCCACGAGGCTTAAATTTCAAGCGCCTTTAAA 862
DB 684 ValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLys 703
QY 863 TATAAATGGTCAAAATAGAAATTTGACGATCAGGTGGAGGACTCCAATATCTAGCTTCT 922
DB 704 TyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyrLeuAlaSer 723
QY 923 CGATATGATTTTCATTGATTCATGATCGTGTGGGCATCCAGGCTGCTCTATGAGGATAC 982
DB 724 ArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyrGlyGlyTyr 743
QY 983 CTCCTCCCTGATGGCATTAAATGCAGAGGTACAGATATCTTCAGGGTTGCTATTGCTGGGCG 1042
DB 744 LeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAlaGlyAla 763
QY 1043 CCAGTCTCTGTGGATCTTCTATGATACAGGATACAG 1081
DB 764 ProValThrLeuTrpIlePheTyrAspThrGlyTyrThr 776

RESULT 5

US-10-311-035-9
; Sequence 9, Application US/10311035
; Publication No. US20040023243A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: YUE, Henry
; APPLICANT: ELLIOTT, Vicki
; APPLICANT: GANDHI, Ameena R.
; APPLICANT: LAL, Preeti
; APPLICANT: AU-YOUNG, Janice
; APPLICANT: TRIBOULEY, Catherine M.
; APPLICANT: DELEGEANE, Angelo M.
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: NGUYEN, Dannie B.
; APPLICANT: LEE, Ernestine A.
; APPLICANT: HAFALIA, April
; APPLICANT: KHAN, Farrah A.
; APPLICANT: CHAWLA, Narinder K.
; APPLICANT: YAO, Monique G.
; APPLICANT: LU, Dyoung Aina M.
; APPLICANT: ARVIZU, Chandra S.
; APPLICANT: TANG, Y. Tom
; APPLICANT: WALSH, Roderick T.
; APPLICANT: AZIMZAI, Valda
; APPLICANT: LU, Yan
; APPLICANT: RAMKUMAR, Jayalaxmi
; APPLICANT: REDDY, Roopa
; APPLICANT: DAS, Depopriya
; APPLICANT: KEARNEY, Liam
; APPLICANT: KALLICK, Deborah A.
; TITLE OF INVENTION: Proteases
; FILE REFERENCE: PI-0123 PCT
; CURRENT APPLICATION NUMBER: US/10/311,035
; CURRENT FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: 60/212,336; 60/213,955; 60/215,396; 60/216,821; 60/218,946
; PRIOR FILING DATE: 2000-06-16; 2000-06-22; 2000-06-29; 2000-07-07; 2000-07-14

```
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PERL Program
; SEQ ID NO 9
; LENGTH: 882
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20040023243A1 7160544CD1
US-10-311-035-9
```

```
Alignment Scores:
Pred. No.: 9 51e-182 Length: 882
Score: 1836.50 Matches: 360
Percent Similarity: 67.5% Conservative: 0
Best Local Similarity: 67.5% Mismatches: 0
Query Match: 94.8% Indels: 173
DB: 4 Gaps: 1
```

```
US-10-825-632-8 (1-1083) x US-10-311-035-9 (1-882)
```

```
QY 2 GAAGAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAGAAGAATTGTGATAGA 61
Db 244 GluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGluGluPheAspArg 263
QY 62 TATCTGGCTATTGTGGTGTCCAAAGCTCAAACTCCCAAGTGGTGTGTAATAATCTT 121
Db 264 TyrSerGlyTyrTrpCysProLysAlaGluThrProSerGlyGlyLysLeu 283
QY 122 AGAATCTATATGAAGAAATCATGTAATCTGAGGTGGAAATTTATCATGTTACATCCCT 181
Db 284 ArgGluLeuTyrGluGluAsnAspGluSerGluValGluLeuHisValThrSerPro 303
QY 182 ATGTGGAAACAGAGGCGAGATTCATTCGTTATCTCTAAACAGTACACCAATCCT 241
Db 304 MetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThrAlaAsnPro 323
QY 242 AAAGTCACCTTTAAGATGTCAGAAATATGATGATGCTGAAGRAGGATCATATA 295
Db 324 LysValThrPheLysMetSerGluLeuMetIleAspAlaGluGlyArgIleIleAspVal 343
QY 295 ----- 295
Db 344 IleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGluTyrIleAla 363
QY 295 ----- 295
Db 364 ArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuAspArgSerGln 383
QY 295 ----- 295
Db 384 ThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGluAspVal 403
QY 295 ----- 295
Db 404 MetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeuIleIleTyr 423
QY 295 ----- 295
Db 424 GluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPheProGlnSer 443
QY 295 ----- 295
Db 444 HisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPheArgHisLeu 463
QY 295 ----- 295
Db 464 TyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGlyGlyLeuPro 483
QY 295 ----- 295
Db 484 AlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSerGlyGluTrp 503
```

```
QY 296 -----GTTGATCAAGTCCAGAGCGTGGTATAT 322
Db 504 GluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArgLeuValTyr 523
QY 323 TTTGAAGCGCAAAAGACTCCCTTTAGAGCATCACCTGTAGCTAGTACGTTACGTTAAAT 382
Db 524 PheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSerTyrValAsn 543
QY 383 CCTGGAGAGGTGACAAAGGCTGACTGACCGTGGCTACTCACATTCCTTGTGTCAGTCAG 442
Db 544 ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysAlaSerGln 563
QY 443 CACTGTGACTTCTTTATAAGTATAGTAACTACCAAGAAATCCACACTGTGTCCTT 502
Db 564 HisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeu 583
QY 503 TACAAGCTATCAAGTCTCAAGATGACCCCACTGCAAAACAAAGAGATTTTGGGCCACC 562
Db 584 TyrLysLeuSerSerProGluAspProThrCysLysThrLysGluPheTrpAlaThr 603
QY 563 ATTTTGGATTTCAGCAGGTCTCTCTGACTATATCTCTCCAGAAATTTTCTTTTGA 622
Db 604 IleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePheSerPheGlu 623
QY 623 AGTACTACTGGATTTACATTTGATGGATGCTCTACAGCCTCATGATCTACAGCCTGGA 682
Db 624 SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly 643
QY 683 AAGAATATCTTACTGTGCTGTTTCATATATGCTGCTCTCAGGTGCTCTCAGGTGCAATTAAT 742
Db 644 LysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeuValAsnAsn 663
QY 743 CGGTTTAAAGAGTCAAGTATTTCCGTTTGAATACCTTCAATACCTCTAGCTCTTAGTGTGTT 802
Db 664 ArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrValVal 683
QY 803 GTAGTATAGACAAACAGGGGATCTCTGTCACGAGGGCTTAAATTTGAAGCGCTTTAAA 862
Db 684 ValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLys 703
QY 863 TATAAAATGGGTCAAATAGAAATTCAGCATCAGGTGGAAGGACTCCAATATCTAGCTTCT 922
Db 704 TyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyrLeuAlaSer 723
QY 923 CGATATGATTTTCATTGACTTAGATCGTGGCATCCAGCGCTGGTCCCTATGAGGATAC 982
Db 724 ArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyrGlyGlyTyr 743
QY 983 CTCTCCCTGATGGCATTAAATGCAGAGTCAAGATATCTTCAGGGTTCCTATTCTGGGCC 1042
Db 744 LeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAlaGlyAla 763
QY 1043 CCAGTCACTCTGTGATCTTCTATGATACAGGATACAG 1081
Db 764 ProValThrLeuTrpIlePheTyrAspThrGlyTyrThr 776
```

RESULT 6

```
US-10-072-012-622
; Sequence 622 Application US/10072012
; Publication No. US20040033493A1
; GENERAL INFORMATION:
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zernhusen, Bryan
; APPLICANT: Patturajan, Meera
; APPLICANT: Shimkets, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangolli, Beha
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie
```

```
; APPLICANT: Taupier Jr, Raymond J.
; APPLICANT: Chusev, Vladimir Y.
; APPLICANT: Colman, Steven D.
; APPLICANT: Molenc, Adam R.
; APPLICANT: Pena, Carol E. A
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Grosse, William M.
; APPLICANT: Alsobrook II, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-258
; CURRENT APPLICATION NUMBER: US/10/072,012
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/265,102
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265,514
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/266,406
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/266,767
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/267,057
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/266,975
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/267,459
; PRIOR FILING DATE: 2001-02-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1391
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 622
; LENGTH: 882
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-072-012-622

Alignment Scores:
Pred. No.: 9,51e-182 Length: 882
Score: 1836.50 Matches: 360
Percent Similarity: 67.5% Conservative: 0
Best Local Similarity: 67.5% Mismatches: 0
Query Match: 94.8% Indels: 173
DB: 4 Gaps: 1

US-10-825-632-8 (1-1083) x US-10-072-012-622 (1-882)
QY 2 GAAGAAGATCCAGATCAGCTGGAGTCGTACTTTGTTCTCCAGAAGAATTGTGATAGA 61
DB 244 GluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluPheAspArg 263
QY 62 TATTCGGCTATTGGTGGTGTCCAAAAGCTGAAACACTCCAGTGGTGTAAATCTTT 121
DB 264 TyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGlyLysIleLeu 283
QY 122 AGAATCTATATGAAGAAATCATGAATCTGAGGTGGAAATATTTCATGTTTACATCCCT 181
DB 284 ArgIleuTyrGluGluAsnAspGluSerGluValGluIleIleHisValThrSerPro 303
QY 182 ATGTTGGAACACAGGAGGCGAGATTCAATCCGTTATCCATAAACAGGTACAGCAATCCT 241
DB 304 MetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThrAlaAsnPro 323
QY 242 AAGTCACTTTTAAGATGTCAGAAATAATGATTGATGCTGAAGGAGGATCATATA 295
DB 324 LysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIleIleAspVal 343
```

```
QY 295 ----- 295
DB 344 IleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGluTyrIleAla 363
QY 295 ----- 295
DB 364 ArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAspArgSerGln 383
QY 295 ----- 295
DB 384 ThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGluAspVal 403
QY 295 ----- 295
DB 404 MetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeuIleIleTyr 423
QY 295 ----- 295
DB 424 GluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPheProGlnSer 443
QY 295 ----- 295
DB 444 HisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPheArgHisLeu 463
QY 295 ----- 295
DB 464 TyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGlyGlyLeuPro 483
QY 295 ----- 295
DB 484 AlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSerGlyGluTrp 503
QY 296 ----- GTTGATGAAGTCAGAACGCTGGTATAT 322
DB 504 GluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArgLeuValTyr 523
QY 323 TTTGAGGACCAACAGACTCCCTTTAGAGATCATCCCTGTACTAGTGTAGTTCAGTTAAAT 382
DB 524 PheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSerTyrValAsn 543
QY 383 CCTGGAGAGGTGACAAAGGCTGACTGACCGTGGCTACTCACATTTCTGTCAGTCAG 442
DB 544 ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCysIleSerGln 563
QY 443 CACTGTGACTCTTTTATAAGTATAGTATAACCAAGAACATCCACACTGTGTGCTCCTT 502
DB 564 HisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeu 583
QY 503 TACAAGCTATCAAGTCCTGAAAGATGACCCAACTTGCAAAACAAAGAAATTTTGGGCCACC 562
DB 584 TyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPheTrpAlaThr 603
QY 563 ATTTGTGATTCAGAGGTCCTCTTCTGACTATCTCTCCAGAAATTTCTCTTTGAA 622
DB 604 IleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePheSerPheGlu 623
QY 623 AGTACTACTCGATTTTACATTGTATGGATGCTCTACAAGCCTCATGATCTACAGCCTGGA 682
DB 624 SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly 643
QY 683 AAGAAATATCCTACTGTGCTGTTTCATATATGTTGGTCTCTCAGTGCAGTGGTGTGATTAAT 742
DB 644 LysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeuValAsnAsn 663
QY 743 CGTTTAAAGAGCTCAAGTATTTCCGCTTGAATACCTAGCTCTCTAGGTTATGTGTT 802
DB 664 ArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrValVal 683
QY 803 GTAGTGATAGACACAGGGGATCCTGTCCACCGAGGCTTAAATTTGAAGGCGCCTTAAA 862
DB 684 ValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLys 703
```

QY 863 TATATAATGGTCAAAATAGAAATTGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCT 922
Db |||||||
QY 704 TyrLysMetGlyGlnIleGluLeuAspGlnValGluGlyLeuGlnTyrLeuAlaSer 723
Db |||||||
QY 923 CGATATGATTTTACATTTAGATCGTGTGGGCATCCAGCGCTGCTCATGGAGGATAC 982
Db |||||||
QY 724 ArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrSerTyrGlyGlyTyr 743
QY 983 CTCTCCCTGATGCTAATTAATGAGAGGTACAGATCTTCACGGTGTCTATTGCTGGGGCC 1042
Db |||||||
QY 744 LeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAlaGlyAla 763
QY 1043 CCAGTCACTGTGGGATCTTCTATGATACAGATACAGC 1081
Db |||||||
QY 764 ProValThrLeuTrpIlePheTyrAspThrGlyTyrThr 776

RESULT 7

US-10-415-122-6
; Sequence 6, Application US/10415122
; Publication No. US20040053369A1
; GENERAL INFORMATION:
; APPLICANT: THE UNIVERSITY OF SYDNEY
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES
; FILE REFERENCE: FPI5217
; CURRENT APPLICATION NUMBER: US/10/415.122
; CURRENT FILING DATE: 2003-08-07
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 882
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-415-122-6

Alignment Scores:

Pred. No.: 9,51e-182 Length: 882
Score: 1836.50 Matches: 360
Percent Similarity: 67.5% Conservative: 0
Best Local Similarity: 67.5% Mismatches: 0
Query Match: 94.8% Indels: 173
DB: 4 Gaps: 1

US-10-825-632-8 (1-1083) x US-10-415-122-6 (1-882)

QY 2 GAAGAAGATGCCAGATCAGCTGGAGTCGCTACCTTGTCTCCAAAGAAATTTGATAGA 61
Db |||||||
QY 62 TATCTGGCTATTGGTGGTCTCCAAAGCTGAACACTCCAGTGGTGGTAAATCTT 121
Db |||||||
QY 264 TyrSerGlyTyrTrpTrpProLysAlaGluThrThrProSerGlyGlyLysIleLeu 283
QY 122 AGAATCTATATGAGAAATGATGATCTGAGGTGGAAATPATTCATGTTACATCCCT 181
Db |||||||
QY 284 ArgIleLeuTyrGluGluLeuAspGluSerGluValGluIleHisValThrSerPro 303
QY 182 ATGTTGGAACAAGGAGGCGAGATTCATTCCTGTTATCTTAAACAGAGGTACAGCAATCCT 241
Db |||||||
QY 304 MetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThrAlaAsnPro 323
QY 242 AAGTCACTTTTAAGATGTCAGAAATTAATGATTGATGCTGAAGGAAGGATCATATA 295
Db |||||||
QY 324 LysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIleIleAspVal 343
QY 295 ----- 295
Db 344 IleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGluTyrIleAla 363
QY 295 ----- 295
Db 364 ArgAlaGlyTyrTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAspArgSerGln 383
QY 295 ----- 295

Db 384 ThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGluAspVal 403
QY 295 ----- 295
Db 404 MetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeuIleIleTyr 423
QY 295 ----- 295
Db 424 GluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPheProGlnSer 443
QY 295 ----- 295
Db 444 HisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPheArgHisLeu 463
QY 295 ----- 295
Db 464 TyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGlyGlyLeuPro 483
QY 295 ----- 295
Db 484 AlaProSerAspPheLysCysProIleLysGluLeuAlaIleThrSerGlyGluTrp 503
QY 296 ----- 322
Db 504 GluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArgLeuValTyr 523
QY 323 TTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTACGTAGCTAGCTTACGTAAT 382
Db |||||||
QY 524 PheGluGlyThrLysAspSerProLeuGluHisLysLeuTyrValValSerTyrValAsn 543
QY 383 CTGGAGAGGTGACAAGGCTCACTGACGCTGCTGCTACATCTTCTGCTCAGCTCAG 442
Db |||||||
QY 544 ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCysIleSerGln 563
QY 443 CACTGTGATCTTTTATTAAGTAAGTATAGTAACAGAGAATCCACACTGTGTGCTCCCT 502
Db |||||||
QY 564 HisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeu 583
QY 503 TACAAGCTATCAAGTCTCTGAAGATGACCCAACTTGGCAAAACAAGAAATTTGGGCCACC 562
Db |||||||
QY 584 TyrLysLeuSerSerProGluAspProThrCysLysThrLysGluPheTrpAlaThr 603
QY 563 ATTTGGGATTCAGCAGGTCTCTTCTGATATATCTCTCCAGAAATTTCTCTTTGAA 622
Db |||||||
QY 604 IleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePheSerPheGlu 623
QY 623 AGTACTACTGGATTTACATTTGATGGATGCTCTACAAGCTCATGATCTACAGCTGGA 682
Db |||||||
QY 624 SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly 643
QY 683 AAGAAATATCTCTGCTGCTGCTCATATATGTTGGTCTCTCAGTGCAGTTGGTGAATAAT 742
Db |||||||
QY 644 LysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeuValAsnAsn 663
QY 743 CGGTTAAAGGAGTCAAGTATTTCCCTTGAATACCTTAGCTCTCTAGGTTATGTGTT 802
Db |||||||
QY 664 ArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrValVal 683
QY 803 GTAGTGATACACACAGGGGATCTCTGTCACGAGGCTTAAATTTGAAGGCCCTTTAAA 862
Db |||||||
QY 684 ValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLys 703
QY 863 TATATAATGGTCAATAGAAATTCAGCATCAGGTGGAAGGACTCCAATATCTAGCTTCT 922
Db |||||||
QY 704 TyrLysMetGlyGlnIleGluLeuAspGlnValGluGlyLeuGlnTyrLeuAlaSer 723
QY 923 CGATATGATTTTCACTTAGATCTGTGGGCATCCAGCGCTGCTCTATGGAGGATAC 982
Db |||||||
QY 724 ArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrTrpSerTyrGlyTyr 743
QY 983 CTCTCCCTGATGCAATTAATGCAGAGGTGAGATATCTTCAAGGTTGCTTATGCTGGGCC 1042
Db |||||||

```

Db      744 LeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAlaGlyAla 763
QY      1043 CCAGTCACCTCTGTGATCTTCTATCATACAGATACAG 1081
Db      764 ProValThrLeuTrpIlePheTyrAspThrGlyTyrThr 776

RESULT 8
US-10-825-632-1
; Sequence 1, Application US/10825632
; Publication No. US20040191826A1
; GENERAL INFORMATION:
; APPLICANT: ABBOTT, Catherine Anne
; APPLICANT: GORRELL, Mark Douglas
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES
; FILE REFERENCE: FCSB-100-Div. 1
; CURRENT APPLICATION NUMBER: US/10/825,632
; CURRENT FILING DATE: 2004-04-15
; PRIOR APPLICATION NUMBER: US 10/070,464
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: PCT/AU00/01085
; PRIOR FILING DATE: 2000-09-11
; PRIOR APPLICATION NUMBER: AU P05709
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: AU P02762
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 882
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-825-632-1

Alignment Scores:
Pred. No.:          9,51e-182      Length:          882
Score:             1836.50      Matches:         360
Percent Similarity: 67.5%      Conservative:    0
Best Local Similarity: 67.5%   Mismatches:     0
Query Match:       94.8%      Indels:         173
DB:                  4        Gaps:              1

US-10-825-632-8 (1-1083) x US-10-825-632-1 (1-882)
QY      2 GAAGAAGATGCCAGATCAGTCGAGTCGCTACCTTTGTTCTCCAAAGAAATTTGATAGA 61
Db      244 GluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGluPheAspArg 263
QY      62 TATTCGGCTATTGGTGTGTCCTCAAAAGCTGAAACAACTCCAGTGTGTGTAATAATCTT 121
Db      264 TyrSerGlyTyrTrpCysProLysAlaGluThrThrProSerGlyGlyLysIleLeu 283
QY      122 AGAATTCATATGAGAAATATGATGAATCTCGAGTGGAAATTTATTCATGTTACATCCCT 181
Db      284 ArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisValThrSerPro 303
QY      182 ATCTTGGAACCAAGGAGGCGAGATTCATTCGTTATCTTAAACAGGTACAGCAAAATCCT 241
Db      304 MetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThrAlaAsnPro 323
QY      242 AAGTCACATTTTAAGATGTCAGAAATATGATTGATGCTGAAGGAGGATCATATA----- 295
Db      324 LysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIleIleAspVal 343
QY      295 ----- 295
Db      344 IleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGluTyrIleAla 363
QY      295 ----- 295
Db      364 ArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAspArgSerGln 383
QY      295 ----- 295

384 ThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGluAspAspVal 403
295 ----- 295
404 MetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeuIleIleTyr 423
295 ----- 295
424 GluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPheProGlnSer 443
295 ----- 295
444 HisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPheArgHisLeu 463
295 ----- 295
464 TyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGlyGlyLeuPro 483
295 ----- 295
484 AlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSerGlyGluTrp 503
296 ----- GTTGATGAAGTCAGAGGCTGTGATAT 322
504 GluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArgLeuValTyr 523
323 TTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTAGTAGTCAGTTACGTAAT 382
524 PheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSerTyrValAsn 543
383 CCTGAGAGGTGACAAAGCTGACTGACCGTGGCTACTCACATTCTTGTGTCATCAGTCAG 442
544 ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCysIleSerGln 563
443 CACTGTGACTCTTTTATAAGTATAGTATAGTAACACAGAGAATCCACACTGTGTGTCCCTT 502
564 HisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeu 583
503 TACAAGCTATCAAGTCTCGAAGATGACCACTTGGCAACAAAGAAATTTGGGGCCACC 562
584 TyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPheTrpAlaThr 603
563 ATTTTGGATTTCAGAGGTCCTCTTCTGACTATATCTCTCCAGAAATTTTCTCTTTTGA 622
604 IleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePheSerPheGlu 623
623 AGTACTACTGGATTTTACATTTGTATGGGATGCTCTACAGCCTCATGATCTACAGCCTGA 682
624 SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly 643
683 AAGAAATATCTTACTGTGTCTTATATATATGTTGGTCTCTCAGTGCAGTTGGTGAATAAT 742
644 LysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeuValAsnAsn 663
743 CGGTTTAAAGGATCAAGTATTTTCGCTTGAATACCTTAGCCTCTCTAGGTTATGTGTT 802
664 ArgSheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrValVal 683
803 GTAGTGATGATACACAAAGGGATCTCTGTCACCGAGGCTTAAATTTGAAGGCGCTTTAAA 862
684 ValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLys 703
863 TATTAATGGTCAAAATAGAAATTCAGATCAGGTGGAGGACTCCATATATCTAGCTTCT 922
704 TyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyrLeuAlaSer 723
923 CGATATCATTTTCATTTGACTTTAGATCTGTGGGATCCACGGCTGTCTATGGAGGATAC 982
724 ArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyrGlyGlyTyr 743
983 CTCTCCCTGATGGCAATTAATGCAGAGTCAGATATCTTCAGGGTGTCTATTCTGGGGCC 1042
744 LeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAlaGlyAla 763

```

QY 1043 CCAGTCTCTGGTGGTCTTCTATGATACAGGATACACG 1081
Db 764 ProValThrLeuTrpPheThrAspThrGlyTyrThr 776
RESULT 9
US-10-982-512-1
; Sequence 1, Application US/10982512
; Publication No, US2005059081A1
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/10/982,512
; CURRENT FILING DATE: 2004-11-05
; PRIOR APPLICATION NUMBER: US/09/976,674
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 882
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-982-512-1

Alignment Scores:
Pred. No.: 9 51e-182 Length: 882
Score: 1836.50 Matches: 360
Percent Similarity: 67.5% Conservative: 0
Best Local Similarity: 67.5% Mismatches: 0
Query Match: 94.8% Indels: 173
DB: 5 Gaps: 1

US-10-825-632-8 (1-1083) x US-10-982-512-1 (1-882)

QY 2 GAAGAAGTCCAGATCAGCTGGAGTCGGTACCTTTGTTCTCCAAAGAAATTTGATAGA 61
Db 244 GluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluPheAspArg 263
QY 62 TATTCTGGCTATTGGTGGTGTCCAAAGCTGAAACAACTCCAGTGGTGGTAAATTTCTT 121
Db 264 TyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGlyLysIleLeu 283
QY 122 AGAATTCATATGAAGAAATGATGAATCTGAGGTGGAAATTTATTGTTTACATCCCT 181
Db 284 ArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisValThrSerPro 303
QY 182 ATGTTGAAACAAGAGGCGCAGATTCAATCCGTTATCTTAAACAGGTACAGCAATCCT 241
Db 304 MetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThrAlaAsnPro 323
QY 242 AAAGTCACCTTTAAGATGTCAAAATAATGATGCTGAGGAGGATCATA----- 295
Db 324 LysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIleIleAspVal 343
QY 295 ----- 295
Db 344 IleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGluTyrIleAla 363
QY 295 ----- 295
Db 364 ArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAspArgSerGln 383
QY 295 ----- 295
Db 384 ThrArgLeuGlnIleValIleSerProGluLeuPheIleProValGluAspVal 403
QY 295 ----- 295

Db 404 MetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeuIleIleTyr 423
QY 295 ----- 295
Db 424 GluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPheProGlnSer 443
QY 295 ----- 295
Db 444 HisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPheArgHisLeu 463
QY 295 ----- 295
Db 464 TyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGlyGlyLeuPro 483
QY 295 ----- 295
Db 484 AlaProSerAspPheLysCysProIleLysGluIleAlaIleThrSerGlyGluTrp 503
QY 296 -----GTTGATGAAGTCAGAGCGCTGTATAT 322
Db 504 GluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgLeuValTyr 523
QY 323 TTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTACGTACGTACGTACGTAAAT 382
Db 524 PheGluGlyThrLysAspSerProLeuGluHisIleLeuTyrValValSerTyrValAsn 543
QY 383 CCTGGAGAGGTGACAGGCTGACTGACCGTGGCTACTCACATTTCTTGCTGCATCAGTAC 442
Db 544 ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCysIleSerGln 563
QY 443 CACTGTGACTCTTTTATAAGTAAGTATAGTAACAGAGAATCCACACTGTGTGTCCCTT 502
Db 564 HisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeu 583
QY 503 TACAAGCTATCAAGTCTCTGAAGATGACCCCAACTGCAAAACAAAGAAATTTGGGCGCAC 562
Db 584 TyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPheTrpAlaThr 603
QY 563 ATTTGGATTCAGAGGTCTCTCTGCTGCTTCAATATCCCTAGCTCTCTAGTGTATGTGTT 622
Db 604 IleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePheSerPheGlu 623
QY 623 AGTACTACTGGATTTACATTTGATGGGATCTCTACAAGCCTCATGATCTACAGCTCGA 682
Db 624 SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly 643
QY 683 AAGAAATATCTACTGTGCTGTTTCATATAAGTGGTCTCTCAGTGCAGTTCGTGAATAAT 742
Db 644 LysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeuValAsnAsn 663
QY 743 CGGTTTAAAGGAGTCAAGTATTTCCGCTTCAATATCCCTAGCTCTCTAGTGTATGTGTT 802
Db 664 ArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrValVal 683
QY 803 GTAGTATAGACAAACAGGGGATCTCTGTCACCGAGGCTTAAATTTGAAGGCGCTTTAAA 862
Db 684 ValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLys 703
QY 863 TATAAAATGGTCAATAGAAATTCACCATCAGGTGAGGAGTCCCAATATCTAGCTTCT 922
Db 704 TyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyrLeuAlaSer 723
QY 923 CGATATGATTTTCATTTAGCTGCTGCGCATCCACCGCTCCATGAGGAGTATC 982
Db 724 ArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrSerTyrGlyTyr 743
QY 983 CTCCTCCCTGATGCTTAATTCAGAGGTTCAGATATCTTTCAGGGTTCGTATTCGTGGGCC 1042
Db 744 LeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAlaGlyAla 763
QY 1043 CCAGTCACTCTGTGGATCTTCTATGATACAGGATACACG 1081
|||||


```
Db 605 ILEUASPSEALAGLYPROLEUPROAEPYRTHRPROGLULILEPHESEPHGLU 624
QY 623 AGTACTACTGGATTTACATTGATGGATGCTCTCAAGCCTCATGATCTACAGCCTGGA 682
Db 625 SerThrThrGlyPheThrLeuYrGlyMetLeuYrLysProHisAspLeuGlnProGly 644
QY 683 AAGAAATATCTACTGTGTGTTTCATATATGCTGCTCAGGTGCAGTGGTGAATAAT 742
Db 645 LysLysTyProThrValLeuPheLeuYrGlyGlyProGlnValGlnLeuValAsnAsn 664
QY 743 CGGTTTAAAGGAGTCAAGTATTTCCGCTTGATACCCCTAGCCTCTCTAGTGTAGGTT 802
Db 665 ArgPheLysGlyValLysTyPheArgLeuAsnThrLeuAlaSerLeuGlyTyValVal 684
QY 803 GTAGTCATAGACAAACAGGGATCCCTGTCACCGAGGGCTTAAATTTGAAGCGCCTTAAA 862
Db 685 ValValLileAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLys 704
QY 863 TATAAATGGTCAATAGAAATTCAGCATCAGGTGGAAGGACTCCAATATCTAGCTTCT 922
Db 705 TyLysMetGlyGlnLileGluLileAspAspGlnValGluGlyLeuGlnTyLeuAlaSer 724
QY 923 CGATATGATTTCAATGACTTAGCTGCTGCTGGCATCCAGCGCTGGTCCCTATGGAGGATAC 982
Db 725 GlnTyAspPheLileAspLeuAspArgValGlyLileHisGlyTrpSerTyrglyGlyTy 744
QY 983 CTCCTCCCTGAGGCATTAATCGAGGTCAGATATCTTCAGGGTTCCTATTGCTGGGCC 1042
Db 745 LeuSerLeuMetAlaLeuMetGlnArgSerAspLilePheArgValAlaLileAlaGlyAla 764
QY 1043 CAGTCATCTGTGGATCTTCTTATGATACAGGATACAGC 1081
Db 765 ProValThrLeuTrpLilePheTyAspThrGlyTyThr 777

RESULT 11
US-10-275-505-2
; Sequence, 2, Application US/10275505
; Publication No. US20040081961A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: DELEGANE, Angelo M.; LAL, Preeti G.
; APPLICANT: HAPALIA, April J.A.; PATTERSON, Chandra
; APPLICANT: WALIA, Narinder K.; KEARNEY, Liam
; APPLICANT: TRIBOULEY, Catherine M.; KHAN, Farrah A.
; APPLICANT: YAO, Monique G.; BAUGHN, Mariah R.
; APPLICANT: AZIMZAI, Yalda; ELLIOTT, Vicki S.
; APPLICANT: NGUYEN, Daniel B.; GANDHI, Ameena R.
; APPLICANT: YANG, Junming; HERNANDEZ, Roberto
; APPLICANT: POLICKY, Jennifer L.; LU, Dyung Aina M.
; APPLICANT: REDDY, Roopa M.; YUE, Henry
; APPLICANT: TANG, Y. Tom
; TITLE OF INVENTION: PROTEASES
; FILE REFERENCE: PI-0085 USN
; CURRENT APPLICATION NUMBER: US/10/275,505
; CURRENT FILING DATE: 2002-11-04
; PRIOR APPLICATION NUMBER: PCT/US01/14651
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: 60/209,402
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: 60/207,477
; PRIOR FILING DATE: 2000-05-24
; PRIOR APPLICATION NUMBER: 60/205,803
; PRIOR FILING DATE: 2000-05-17
; PRIOR APPLICATION NUMBER: 60/203,566
; PRIOR FILING DATE: 2000-05-11
; PRIOR APPLICATION NUMBER: 60/202,082
; PRIOR FILING DATE: 2000-05-04
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PERL Program
; SEQ ID NO 2
; LENGTH: 580
; TYPE: PRT
```

```
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 376067CD1
US-10-275-505-2
```

```
Alignment Scores:
Pred. No.: 3,02e-146 Length: 580
Score: 1495.00 Matches: 306
Percent Similarity: 57.8% Conservative: 0
Best Local Similarity: 57.8% Mismatches: 1
Query Match: 77.1% Indels: 223
DB: 4 Gaps: 2
```

US-10-825-632-8 (1-1083) x US-10-275-505-2 (1-580)

```
QY 9 ATCCGAGATCAGCTGGAGTCGCTACCTTGTCTCTCCAGAAGAATTTGATAGATATTCTG 68
Db 1 MetProAspGlnLeuGluSerLeuProLeuPheSerLysLysAsnLeuLileAspLileLeu 20
QY 69 GCTATTGGTGGTGTCCAAAGCTGAAACAACTCCCACTGGTGGTAAATTTCTTAGAATTC 128
Db 21 AlaLileGlyValGlnLysLeuLysGlnLeuProValValLysPheLeuGluPhe 40
QY 129 TATATGAGAAATATGATCAATCTCAGGTGGAAATTTATTCATGTTCATCCCTATGTGG 188
Db 41 TyrMetLysLysMetMetAsnLeuArgTrpLysLeuPheMetLeuHisProLeuCysTrp 60
QY 189 AAACAAGGAGG- GCAGATTTCATTCCTGTTATCTTAAACAGGTACAGCAATCTCTAAGTC 247
Db 61 LysGlnGlyArgAlaAspSerPheArgTyProLysThrGlyThrAlaAsnProLysVal 80
QY 248 ACTTTAAAGATGTCAGAAATATGATTGATGCTGGAAGGAGGATCATA----- 295
Db 81 ThrPheLysMetSerGluLileMetLileAspAlaGluGlyArgLileAspValLileAsp 100
QY 295 ----- 295
Db 101 LysGluLeuLileGlnProPheGluLileLeuPheGluGlyValGluTyRileAlaargala 120
QY 295 ----- 295
Db 121 GlyTrpThrProGluGlyLysTyRileAlaTrpSerLileLeuLeuAspArgSerGlnThrArg 140
QY 295 ----- 295
Db 141 LeuGlnLileValLeuLileSerProGluLeuPheLileProValGluAspValMetGlu 160
QY 295 ----- 295
Db 161 ArgGlnArgLeuLileGluSerValProAspSerValThrProLeuLileTyRileGluGlu 180
QY 295 ----- 295
Db 181 ThrThrAspLileTrpLileAsnLileHisAspLilePheHisValPheProGlnSerHisGlu 200
QY 295 ----- 295
Db 201 GluGluLileGluPheLilePheAlaSerGluCysLysThrGlyPheArgHisLeuTyRileLys 220
QY 295 ----- 295
Db 221 IleThrSerLileLeuLysGluSerLysTyRileLysArgSerSerGlyGlyLeuProAlaPro 240
QY 295 ----- 295
Db 241 ThrValThrTrpMetLileThrPheMetArgSerLeuGlyThrProSerCysMetCysVal 260
QY 296 -----GTTGATCAAGTCAAGAGGCTGGTATATTTTGAAGGCACC 334
Db 261 ThrHisLileValGluLileGlnValAspGluValArgArgLeuValTyRileGluGlyThr 280
QY 335 AAAGACTCCCTTTAGAGCATCATCCTGACGTAGTCAGTTACGTAAATCCTCGAGAGGTG 394
```


Db 281 LysAspSerProLeuGluHisLeuTyrValValSerTyrValAsnProGlyGluVal 300
QY 395 ACAGGCTGACTGACGGTGGTACTACATCTTGGCTGCATCAGTACGACTGTGACTTC 454
Db 301 ThrArgLeuThrAspArgGlyTyrSerHisSerCysHisSerGlnHisCysAspPhe 320
QY 455 TTTATAAGTAAGTATAGTAACAGAGAAATCCACACTGTGTGTCCTTTTACAAAGCTATCA 514
Db 321 PheLeuSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeuTyrLysLeuSer 340
QY 515 AGTCCTGAAGATGACCAACTTGCACAAAGAAATTTGGGCCACCATTTGGATTCA 574
Db 341 SerProGluAspAspProThrCysLysThrLysGluPheTrpAlaThrIleLeuAspSer 360
QY 575 GCAGGTCTCTTCTGACTATCTCTCCAGAAATTTCTCTTTTGAAGTACTACTGGA 634
Db 361 AlaGlyProLeuProAspTyrThrProProGluIlePheSerPheGluSerThrThrGly 380
QY 635 TTTACATTGTATGGGATGCTTACAAAGCCTCATGATCTACAGCCTGGAAGAAATATCCT 694
Db 381 PheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGlyLysLysTyrPro 400
QY 695 ACTGTGCTGTTCATATATGTCCTCTCAGGTGCAGTTGGTAATAATCGGTTAAAGGA 754
Db 401 ThrValLeuPheIleTyrGlyGlyProGlnValGlnLeuValAsnAsnArgPheLysGly 420
QY 755 GTCAAGTATTTCGGTTGAATACCTAGCCTCTAGGTATGTTAGTTGTTAGTGATAGAC 814
Db 421 ValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrValValValIleAsp 440
QY 815 ACAGGGGATCTGTCCACGAGCGCTTAAATTTGAAGCGCTTTAAATATAAATGAGT 874
Db 441 AsnArgLysSerCysHisArgGlyLeuLysPheGluGlyAlaPheLysTyrLysMet 459
QY 875 CAATAGAAATTGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTC 934
Db 459 ----- 459
QY 935 ATTGACTTAGATCGTGTGGGCATCCAGCGCTGGTCTTATGAGGATACCTTCCTCGATG 994
Db 459 ----- 459
QY 995 GCATTAAATGACAGGTTCAGATATCTTACGGTTGCTATTGCTGGGGCCCCAGTCACTGT 1054
Db 460 -----ValAlaIleAlaGlyAlaProValThrLeu 469
QY 1055 TGGATCTTCTATGATACAGGATACAG 1081
Db 470 TrpIlePheTyrAspThrGlyTyrThr 478

RESULT 12

US-11-140-224-2
; Sequence 2, Application US/11140224
; Publication No. US2005027280A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: DELEGEANE, Angelo M.; LAL, Preeti G.
; APPLICANT: HAFALIA, April J.A.; PATTERSON, Chandra
; APPLICANT: WALIA, Narinder K.; KEARNEY, Liam
; APPLICANT: TRIBOULEY, Catherine M.; KHAN, Farrah A.
; APPLICANT: YAO, Monique G.; BAUGHN, Mariah R.
; APPLICANT: AZIMZAI, Yalda; ELLIOTT, Vicki S.
; APPLICANT: NGUYEN, Daniel B.; GANDHI, Ameena R.
; APPLICANT: YANG, Junming; HERNANDEZ, Roberto
; APPLICANT: POLICKY, Jennifer L.; LU, Dyung Aina M.
; APPLICANT: REDDY, Roopa M.; YUE, Henry
; APPLICANT: TANG, Y. Tom
; TITLE OF INVENTION: PROTEASES
; FILE REFERENCE: PI-0085 USN
; CURRENT APPLICATION NUMBER: US/11/140,224
; CURRENT FILING DATE: 2005-05-31
; PRIOR APPLICATION NUMBER: US/10/275,505

; PRIOR FILING DATE: 2002-11-04
; PRIOR APPLICATION NUMBER: PCT/US01/14651
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: 60/209,402
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: 60/207,477
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/205,803
; PRIOR FILING DATE: 2000-05-17
; PRIOR APPLICATION NUMBER: 60/203,566
; PRIOR FILING DATE: 2000-05-11
; PRIOR APPLICATION NUMBER: 60/202,082
; PRIOR FILING DATE: 2000-05-04
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PERL Program
; SEQ ID NO 2
; LENGTH: 580
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 376067CD1
US-11-140-224-2
Alignment Scores:
Pred. No.: 3,02e-146 Length: 580
Score: 1495.00 Matches: 306
Percent Similarity: 57.8% Conservatives: 0
Best Local Similarity: 57.8% Mismatches: 1
Query Match: 77.1% Indels: 223
DB: 6 Gaps: 2
US-10-825-632-8 (1-1083) x US-11-140-224-2 (1-580)
QY 9 ATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAAGAAATTTGATAGATATCTG 68
Db 1 MetProAspGlnLeuGluSerLeuProLeuPheSerLysLysAsnLeuIleAspLeu 20
QY 69 GCTATTGCTGCTCCAAAGAGCTGAAACAACTCCCAAGTGGTGGTAAAAATTTCTTAGAATTC 128
Db 21 AlaIleGlyValGlnLysLeuLysGlnLeuProValValLysPheLeuGluPhe 40
QY 129 TATATGAAGAAATGATGAATCTGAGTGTGAAATTTATTCATGTTACATCCCTATGTTGG 188
Db 41 TyrMetLysLysMetMetAsnLeuArgTrpLysLeuPheMetLeuHisProLeuCysTrp 60
QY 189 AAACAAGAGG-GCAGATTTCATTCCTGTTATCTCTAAACAGGTACAGCAATCTCTAAAGTC 247
Db 61 LysGlnGlyArgAlaAspSerPheArgTyrProLysThrGlyThrAlaAsnProLysVal 80
QY 248 ACTTTTAAGATGTCAGAAATTAATGATGTCGTGAAGGAGGATCATTA----- 295
Db 81 ThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIleIleAspValIleAsp 100
QY 295 ----- 295
Db 101 LysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGluTyrIleAlaArgAla 120
QY 295 ----- 295
Db 121 GlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAspArgSerGlnThrArg 140
QY 295 ----- 295
Db 141 LeuGlnIleValLeuIleSerProGluLeuPheIleProValGluAspAspValMetGlu 160
QY 295 ----- 295
Db 161 ArgGlnArgLeuIleGluSerValProAspSerValThrProLeuIleIleTyrGluGlu 180
QY 295 ----- 295
Db 181 ThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPheProGlnSerHisGlu 200

```
QY 295 ----- 295
Db 201 GluGluLeuGluPheLeuPheAlaSerGluCysLysThrGlyPheArgHisLeuTyrLys 220
QY 295 ----- 295
Db 221 IleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGlyLeuProAlaPro 240
QY 295 ----- 295
Db 241 ThrValThrTrpMetIleThrPheMetArgSerLeuGlyThrProSerCysMetCysVal 260
QY 296 ----- 296
Db 261 ThrHisIleValGluLeuValAlaSerGluValArgLeuValTyrPheGluGlyThr 280
QY 335 AAAGACTCCCTTTAGACATACCTGTACGTAGTACGTAGTACGTAACTCTGAGAGGNG 394
Db 281 LysAspSerProLeuGluHisLeuTyrValValSerTyrValAlaSerProGlyGluVal 300
QY 395 ACAAGCTGACTGACCGTGGCTACTCATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 454
Db 301 ThrArgLeuThrAspArgGlyTyrSerHisSerCysCysAlaSerGlnHisCysAspPhe 320
QY 455 TTTATAAGTATAGTATAGTAAACAGAGAATCCACACTGTGTGCTTACAGCTATCA 514
Db 321 PheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeuTyrLysLeuSer 340
QY 515 AGTCTGAAGATGACCCAACTTGCNAACAAAGGAATTTTGGCCACCATTTTGGATTCA 574
Db 341 SerProGluAspAspP-roThrCysLysThrLysGluPheTrpAlaThrIleLeuAspSer 360
QY 575 GCAGGTCTCTTCTGACTATCTCTCCAGAAATTTCTCTTTGAAAGTACTACTGGA 634
Db 361 AlaGlyProLeuProAspTyrThrProProGluIlePheSerPheGluSerThrThrGly 380
QY 635 TTTACATGTATGGATGCTCTACAGCCTCATGATCTACAGCCTCGAAAGAAATATCCT 694
Db 381 PheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGlyLysLysTyrPro 400
QY 695 ACTGTGCTGTATATATGCTGCTCAGTGCAGTGGTGAATATCGTTTAAAGGA 754
Db 401 ThrValLeuPheIleTyrGlyGlyProGlnValGlnLeuValAlaSerPheLysGly 420
QY 755 GTCAAGTATTTCCGCTTGAATACCTAGCCTCTCTAGGTTATGTTGTTAGTAGAGAC 814
Db 421 ValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrValValValIleAsp 440
QY 815 AACAGGGATCTGTACAGGAGGCTTAAATTTGAAGCGCCTTTAAATATATAAATGGGT 874
Db 441 AsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLysTyrLysMet 459
QY 875 CAAATAGAAATGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTC 934
Db 459 ----- 459
QY 935 ATTGACTTAGATCGTGTGGGCATCCACGGCTGGTCTCTATGAGGATACCTCTCCCTGATG 994
Db 459 ----- 459
QY 995 GCATTATGACAGGTGATATCTTTCAGGGTTGCTATTGCTGGGGCCCCAGTCACTCTG 1054
Db 460 ----- 460
QY 1055 TGGATCTTCTATGATACAGGATACAGC 1081
Db 470 TrpIlePheTyrAspThrGlyTyrThr 478
```

RESULT 13

US-09-976-674-7

; Sequence 7, Application US/09976674

; Patent No. US20020115843A1

```
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; PRIOR APPLICATION NUMBER: US 2001-10-12
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 690
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-7
```

Alignment Scores:

Pred. No.:	3,24e-118	Length:	690
Score:	1226.00	Matches:	259
Percent Similarity:	48.5%	Conservative:	0
Best Local Similarity:	48.5%	Mismatches:	0
Query Match:	63.3%	Indels:	275
DB:	3	Gaps:	2

US-10-825-632-8 (1-1083) x US-09-976-674-7 (1-690)

```
QY 2 GAAGAAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAAGAAGATTTGATAGA 61
Db 244 GluGluAspAlaAspSerAlaGlyValAlaThrPheValLeuGlnGluPheAspArg 263
QY 62 TATTCTCGCTATTGGTGGTGTCCAAAGCTGAAACAACTCCAGTGGTGGTAAATTCCT 121
Db 264 TyrSerGlyTyrTrpCysProLysAlaGluThrThrProSerGlyGlyLysIleLeu 283
QY 122 AGAATTCATATCAAGAAATGATGATCTGAGTGGAAATTTATTCATGTTACATCCCT 181
Db 284 ArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleHisValThrSerPro 303
QY 182 ATGTTGGAACAAGAGGCGCAGATTCATTCCTGTTATCTCTAAACAGGTACAGCAATCCT 241
Db 304 MetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThrAlaAsnPro 323
QY 242 AAAGTCACTTTAAGATGTCAGAAAATAATGATGATGCTGAAGGAAGGATCATA----- 295
Db 324 LysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIleIleAspVal 343
QY 295 ----- 295
Db 344 IleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGluTyrIleAla 363
QY 295 ----- 295
Db 364 ArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAspArgSerGln 383
QY 295 ----- 295
Db 384 ThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGluAspAspVal 403
QY 295 ----- 295
Db 404 MetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeuIleIleTyr 423
QY 295 ----- 295
Db 424 GluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPheProGlnSer 443
QY 295 ----- 295
Db 444 HisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPheArgHisLeu 463
```

Mon May 8 13:41:59 2006

TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV

FILE REFERENCE: 70669
CURRENT APPLICATION NUMBER: US/10/982,512
CURRENT FILING DATE: 2004-11-05
PRIOR APPLICATION NUMBER: US/09/976,674
PRIOR FILING DATE: 2001-10-12
PRIOR APPLICATION NUMBER: US 60/240,117
PRIOR FILING DATE: 2000-10-12
NUMBER OF SEQ ID NOS: 61
SOFTWARE: Patentin version 3.1
SEQ ID NO 7
LENGTH: 690
TYPE: PRT
ORGANISM: Homo sapiens
US-10-982-512-7

Alignment Scores: 3.24e-118 Length: 690
Pred. No.: 1226.00 Matches: 259
Score: 1226.00 Conservative: 0
Percent Similarity: 48.5% Mismatches: 0
Best Local Similarity: 48.5% Indels: 275
Query Match: 63.3% Gaps: 2
DB: 5

US-10-825-632-8 (1-1083) x US-10-982-512-7 (1-690)

QY 295 ----- 295
Db 464 TyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGlyGlyLeuPro 483
QY 295 ----- 295
Db 484 AlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSerGlyGluTrp 503
QY 296 ----- 296
Db 504 GluValLeuGlyArgHisGlySerAsnIleGlnValaspGluValArgLeuValTyr 523
QY 323 TTTGAAGCCACCAAGACTCCCTTTAGACCATCCTGTACGTAGTACGTACGTAAAT 382
Db 524 PheGluGlyThrLysAspSerProLeuGluHisLeuIleValValSerTyrValAsn 543
QY 383 CTGGAGAGGTGACAAAGCTGACGCTGGCTACTCACATTTCTGTCATGTCAGTACAG 442
Db 544 ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysIleSerGln 563
QY 443 CACTGTGACTTCTTTATAGTAAGTATAGTAACAGAGAATCCACACTGTGTCCCTT 502
Db 564 HisCysAspPheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeu 583
QY 503 TACAAGCTATCAAGTCTGGAAGATGACCCAACTTGCACAAACAAAGGAATTTTGGGCCACC 562
Db 584 TyrLysLeuSerSerProGluaspAspProThrCysLysThrLysGluPheTrpAlaThr 603
QY 563 ATTTTGGNTACAGAGTCTCTCTCTGACTATATCTCTCCAGAAATTTCTCTTTGAA 622
Db 623 AGTACTACTGATTTACATTTATGGTGTCTTACAGCTCTATGATGATGATGATGATGAT 682
QY 624 SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly 643
Db 683 AAGAAATATCTACTGTGCTGTTTCATATATATGTTGGTCTTACAGTGCAGTGCAGTAAT 742
QY 644 LysLysTyrProThrValLeuPheIleTyrGlyGly----- 655
Db 743 CGGTTTAAAGGAGTCAAGTATTTCCCGTTGAATACCTTAGCTCTCTAGTTATGTGTT 802
QY 655 ----- 655
Db 803 GTAGTGATAGACACAGGGATCTGTCCAGGAGGCTTAAATTTGAAGGCCCTTTAAA 862
QY 655 ----- 655
Db 863 TATAAATGGGTCAAAATAGAAATTGACGATCAGGTGAAGGACTCCAATATCTAGCTTCT 922
QY 655 ----- 655
Db 923 CGATATGATTTCAITGACTTAGATCGTGTGGGCATCCACGGCTGCTCTATGAGGATAC 982
QY 655 ----- 655
Db 983 CTCTCCCTGATGCAATTAATGACAGAGTACAGATATCTTACAGGTGCTATTCTGGGGCC 1042
QY 656 ----- 656
Db 1043 CCAGTCACTGTGGATCTTCTATGATACAGGATACACGG 1082
QY 662 oGlnSerCysGlySerSerMetIleGlnaspThrArg 675
Db

RESULT 14
US-10-982-512-7
Sequence 7, Application US/10982512
Publication No. US20050059081A1
GENERAL INFORMATION:
APPLICANT: Oi, Steve
APPLICANT: Akinsanya, Karen
APPLICANT: Riviere, Pierre
APPLICANT: Junien, Jean-Louis

```
QY 295 ----- 295
Db 484 AlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSerGlyGluTrp 503
QY 296 ----- 296
Db 504 GluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValA-Sa-GluValValTrp 523
QY 323 TTTGAAGGCCCAAGAGCTCCCTTTAGAGCATCACCTGTACGTAGTCAGTCACTGATATAT 322
Db 524 PheGluGlyThrLysAspSerProLeuGluHisLysLeuTrpValValSerTrpValAsn 543
QY 383 CTGTGAGGCTCAAGGCTGACTGACCGTGGCTACTACATCTTCTGCTGCATCAGTCAG 382
Db 544 ProGlyGluValThrArgLeuThrAspArgGlyTrpSerHisSerCysValSerGln 563
QY 443 CACTCTGACTCTTTTAAAGTATAGTATAGTAACTGCAAGAGAAATCCACACTGTGTCCCTT 502
Db 564 HisCysAspPhePheIleSerLysLeuThrAspArgGlyTrpSerHisSerCysValSerGln 563
QY 503 TACAGCTATCACTCTGAGATGACCCAACTTGCAAAACAAAGGAATTTTGGGCCACC 562
Db 584 TyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPheTrpAlaThr 603
QY 563 ATTTGGATTACAGAGCTCTCTCTGACTATPACTCTCCAGAAATTTCTCTTTTGA 622
Db 604 IleLeuAspSerAlaGlyProLeuProAspTrpThrProGluIlePheSerPheGlu 623
QY 623 AGTACTACTGGATTACATGTGGATGCTCTCAAGCCTCATGATCTACAGCCTGGA 682
Db 624 SerThrThrGlyPheThrLeuTrpGlyMetLeuTrpLysProHisAspLeuGlnProGly 643
QY 683 AAGAAATATCTACTGTGTGTATATATGTTGCTCTCAGGTGCAGTGGTGAATAAT 742
Db 644 LysLysTrpProThrValLeuPheIleTrpGlyGly 655
QY 743 CGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTAGCCTCTTAGGTTATGTT 802
Db 655 ----- 655
QY 803 GTAGTGATAGCAACAGGGGATCCTGTCAACGAGGGCTTAAATTTGAGCGCCTTTAA 862
Db 655 ----- 655
QY 863 TATAAAATGGGTCAAATAGAAATGACGATCAGGTGGAGGACTCCCAATATCTAGCTTCT 922
Db 655 ----- 655
QY 923 CGATATGATTTCAATGACTTAGATCGTGTGGCATCCAGGCTGGTCTATGAGGATAC 982
Db 655 ----- 655
QY 983 CTCTCCCTGATGGCATTAAATGCAGAGTCAAGATATCTTCAGGGTTGCTATTGCTGGGCC 1042
Db 656 ----- 656
QY 1043 CCAGTCACTCTGGGATCTTCTATGATACAGGATACAGG 1082
Db 662 OGInSerLeuGlySerSerMetIleGlnAspThrArg 675

RESULT 15
US-10-072-012-620
; Sequence 620, Application US/10072012
; Publication No. US2004003493A1
; GENERAL INFORMATION:
; APPLICANT: Tchenev, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Patturajan, Meera
; APPLICANT: Shimkete, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangolli, Bsha
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Restelli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Tauber, Valerie
; APPLICANT: Guever, Jodi Raymond J.
; APPLICANT: Coleman, Steven D.
; APPLICANT: Wolenc, Adam E. A.
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Grosse, William M.
; APPLICANT: Alsobrook II, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-258
; CURRENT APPLICATION NUMBER: US/10/072,012
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/265,102
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265,514
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/266,406
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/266,767
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/267,057
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/266,975
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/267,459
; Remaining prior application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1391
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 620
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-072-012-620
Alignment Scores:
Pred. No.: 1,21e-117 Length: 689
Score: 1220.50 Matches: 689
Percent Similarity: 54.3% Conservative: 54
Best Local Similarity: 44.2% Mismatches: 69
Query Match: 63.0% Indels: 175
DB: 4 Gaps: 2
US-10-825-632-8 (1-1083) x US-10-072-012-620 (1-689)
QY 5 GAAGATGCCAGATCAGCTGAGTCGCTACCTTTGTTCTCCAGAAAGATTTGATAGATAT 64
Db 60 AspAspProLysSerAlaGlyValAlaThrPheValIleGlnGluPheAspArgPhe 79
QY 65 TCTGGCTATTGGTGGTGTCCAAAGCTGNAACACTCCAGTGGTGGT---AAAATCTT 121
Db 80 ThrGlyTrpTrpTrpCysProThrAlaSerTrpGluGlySerGluGlyLeuLysThrLeu 99
QY 122 AGAATCTATATGAAGAAATGATGAATCTGAGGTGGAAATTTATCATGTTACATCCCT 181
Db 100 ArgIleLeuTrpGluGluValAspGluSerGluValGluValIleHisValProSerPro 119
QY 182 ATGTTGGAACCAAGGAGGCGAGATTTCATTCGTTTCTTAAACAGGTACAGCAATCTCT 241
```

[illegible]

```

Search completed: May 2, 2006, 03:08:58
Job time : 179.086 secs

```

Qy	800	GTTCTACTGATAGCAACACCGGGNTCCTGCACGAGGCGTAAATTGTGAAGCGCCCTTT	859
Db	500	ValValVallleAspGlyArgglySerCysgluArgGlyLeuargPheGludGlyAlaLeu	519
Qy	860	AAATATAAATGGGTCAAATAGAANAATTGACGATCAGGTGGGAAGCATCCCAATATCTAGCT	919
Db	520	LysasnGlnMetGlyGlnValGluIleGluaspGlnValGluGlyLeuGlnPheValAla	539
Qy	920	TCTCGATATGATTTTCATTGACTTAGATCTGTTGGGCATCCACGGCTGGTGCTCTATGAGAGA	979
Db	540	GlulysTyrglyPheIleAspLeuSerArgValAlalleHisGlyTrpSerTyrglyGly	559
Qy	980	TACTCTCCCTGATGGCATTAAATGCAGAGCTCAGATATCTTCAGGGTTGCTATTGCTGGG	1039
Db	560	PheLeuSerLeuMetGlyLeuIleHisLysProGlnValPheLysValAlalleAlaGly	579
Qy	1040	GCCCCAGCTCACTCTGTGGATCTTCTATGATACAGGATACACG	1081
Db	580	AlaProValThrValTrpMetAlaTyAspThrGlyTyThr	593

THIS PAGE BLANK (USPTO)

Sequence 4, Appli
Sequence 13, Appli
Sequence 2, Appli
Sequence 234, Appli
Sequence 2, Appli
Sequence 6, Appli
Sequence 55, Appli
Sequence 7504, Ap
Sequence 1, Appli
Sequence 1591, Ap
Sequence 9281, Ap
Sequence 1365, Ap
Sequence 16, Appli
Sequence 6, Appli
Sequence 16, Appli
Sequence 16, Appli
Sequence 5396, Ap
Sequence 23, Appli
Sequence 44, Appli
Sequence 44, Appli
Sequence 17, Appli
Sequence 53, Appli
Sequence 18, Appli
Sequence 3, Appli
Sequence 6, Appli
Sequence 5933, Ap
Sequence 31619, A
Sequence 638, App
Sequence 639, App
Sequence 14, Appli
Sequence 14, Appli
Sequence 4, Appli
Sequence 26, Appli
Sequence 8562, Ap
Sequence 18967, A
Sequence 375, App
Sequence 16980, A
Sequence 22602, A
Sequence 16979, A
Sequence 22601, A
Sequence 16978, A

11 US-11-208-288-4
11 US-11-116-939-13
9 US-10-522-789-2
9 US-10-501-035-234
9 US-11-208-288-2
11 US-11-208-288-6
11 US-11-186-284-55
11 US-11-079-463-7504
11 US-11-179-977-1
11 US-11-264-096-1591
11 US-11-079-463-9281
11 US-11-045-004-1565
11 US-11-176-951-16
11 US-11-179-977-6
11 US-11-179-977-16
11 US-11-079-463-5396
11 US-11-151-601-23
11 US-11-288-493-44
9 US-10-204-639-44
24 US-11-176-951-17
11 US-11-113-424-53
9 US-10-909-769-18
11 US/11/062
11 US/11/062
11 US-11-079-463-5993
11 US-11-096-568A-31619
9 US-10-995-561-638
9 US-10-995-561-639
11 US-11-108-519-14
11 US-11-141-554B-14
11 US-11-179-977-4
11 US-11-186-284-26
11 US-11-087-099-8562
11 US-11-188-298-18967
9 US-10-216-161A-375
11 US-11-096-568A-16980
11 US-11-096-568A-22602
11 US-11-096-568A-16979
11 US-11-096-568A-22601
11 US-11-096-568A-16978

738 11
762 11
766 9
766 9
766 11
760 11
760 11
624 11
657 11
349 11
737 11
655 11
24 11
298 11
318 11
360 11
710 11
803 9
24 11
2515 11
849 9
1067 11
1092 11
519 11
475 11
441 9
441 11
441 11
300 11
3063 11
925 11
925 11
816 9
317 11
332 11
361 11
364 11
402 11

315 16.3
315 16.3
315 16.3
315 16.3
315 16.3
309 15.9
315 15.9
10.3
199 10.3
196 10.1
194 10.0
141.5 7.3
127.5 6.6
125 6.4
101.5 5.2
101 5.2
100.5 5.2
96 5.0
96 5.0
94.5 4.9
93 4.8
92 4.7
91.5 4.7
91.5 4.7
91.5 4.7
90 4.6
87.5 4.5
87 4.5
87 4.5
87 4.5
86.5 4.5
86.5 4.5
86 4.4
86 4.4
85.5 4.4
85 4.4
85 4.4
85 4.4
85 4.4
85 4.4

ALIGNMENTS

RESULT 1
US-11-151-601-20
; Sequence 20, Application US/11/151601
; Publication No. US20060003413A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Olandt, Peter J.
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Curtis, Rory A. J.
; APPLICANT: Williamson, Mark
; APPLICANT: Welch, Nadine
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASE, PHOSPHATASE,
; AND PROTEASE FAMILY MEMBERS AND USES THEREOF
; FILE REFERENCE: WPI00-054P1RC10M1D1V
; CURRENT APPLICATION NUMBER: US/11/151,601
; CURRENT FILING DATE: 2005-06-13
; PRIOR APPLICATION NUMBER: US 10/170,789
; PRIOR FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: US 09/797,039
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: PCT/US01/06525
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/186,061
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 09/882,166
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: PCT/US01/19269
; PRIOR FILING DATE: 2001-06-15

GenCore version 5.1.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: May 2, 2006, 02:50:19 ; Search time 3.62006 Seconds
(without alignments)
2721.092 Million cell updates/sec

Title: US-10-825-632-8
Perfect score: 1938
Sequence: 1 ggaagaatgccagatcag.....tatgatacagatacacgga 1083

Scoring table: BIOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 232119 seqs, 45477862 residues
Total number of hits satisfying chosen parameters: 464238

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame_n2p.model -DEV=xlp
-Q=/abes/ABSWEB_spool/US10825632/runat_01052006_110000_3533/app_query.fasta 1
-DB=Published Applications AA New -QFMT=fastan -SUFFIX=n2p.rapbn -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blomsum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -HOST=abs802p
-USER=US10825632 @CGN_1_1_59 @runat_01052006_110000_3533 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications AA New:
1: /SIDSS/prodata/1/pubpaa/US08 NEW PUB pep1.*
2: /SIDSS/prodata/1/pubpaa/US06 NEW PUB pep.*
3: /SIDSS/prodata/1/pubpaa/US07 NEW PUB pep.*
4: /SIDSS/prodata/1/pubpaa/US08 NEW PUB pep.*
5: /SIDSS/prodata/1/pubpaa/US05 NEW PUB pep.*
6: /SIDSS/prodata/1/pubpaa/US09 NEW PUB pep.*
7: /SIDSS/prodata/1/pubpaa/US10 NEW PUB pep.*
8: /SIDSS/prodata/1/pubpaa/US11 NEW PUB pep.*
9: /SIDSS/prodata/1/pubpaa/US11 NEW PUB pep.*
10: /SIDSS/prodata/1/pubpaa/US11 NEW PUB pep.*
11: /SIDSS/prodata/1/pubpaa/US60 NEW PUB pep.*
12: /SIDSS/prodata/1/pubpaa/US60 NEW PUB pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	1836.5	94.8	882	11	US-11-151-601-20 Sequence 20, Appli
2	984	50.8	497	11	US-11-264-096-1594 Sequence 1594, Ap
3	980	50.6	497	11	US-11-264-096-1593 Sequence 1593, Ap
4	383	19.8	627	11	US-11-079-463-7758 Sequence 7758, Ap
5	372	19.2	745	11	US-11-079-463-6408 Sequence 6408, Ap

```
; PRIOR APPLICATION NUMBER: US 60/212,078
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 09/934,406
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: PCT/US01/26052
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/226,740
; PRIOR FILING DATE: 2000-08-21
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 882
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-151-601-20

Alignment Scores:
Pred. No.: 1,816-175 Length: 882
Score: 1836.50 Matches: 360
Percent Similarity: 67.5% Conservative: 0
Best Local Similarity: 67.5% Mismatches: 0
Query Match: 94.8% Indels: 173
DB: 11 Gaps: 1

US-10-825-632-8 (1-1083) x US-11-151-601-20 (1-882)

QY 2 GAAGAAGATCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAGAAGAATTGTAGTA 61
Db 244 GluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluPheAspArg 263
QY 62 TATCTGCTGATTTGGTGGTGTCCAAAAGCTGAAACAACCTCCCAAGTGGTGGTAAATCTT 121
Db 264 TyrSerGlyTyrTrpCysProLysAlaGluThrThrProSerGlyGlyLysLeu 283
QY 122 AGAATCTATATGAAGAAATGATGAATCTGAGTGGAAATATTTCATCTACATCCCT 181
Db 284 ArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisValThrSerPro 303
QY 182 ATGTTGGAACAAAGAGGAGGAGATTTCATTCCTGATTCCTAAACAGGTACAGCAATCT 241
Db 304 MetLeuGluThrAlaGalaAspSerPheArgTyrProLysThrGlyThrAlaAsnPro 323
QY 242 AAAGTCACCTTTTAAGATGTCAGAAATAATGATGCTGCTGAAGGAGGATCATA 295
Db 324 LysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIleIleAspVal 343
QY 295 ----- 295
Db 344 IleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGluTyrIleAla 363
QY 295 ----- 295
Db 364 ArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAspArgSerGln 383
QY 295 ----- 295
Db 384 ThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGluAspVal 403
QY 295 ----- 295
Db 404 MetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeuIleIleTyr 423
QY 295 ----- 295
Db 424 GluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPheProGlnSer 443
QY 295 ----- 295
Db 444 HisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPheArgHisLeu 463
QY 295 ----- 295
```

```
Db 464 TyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGlyGlyLeuPro 483
QY 295 ----- 295
Db 484 AlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSerGlyGluTrp 503
QY 296 ----- 296
Db 504 GluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArgLeuValTyr 523
QY 323 TTTGAAGGACCAAGACTCCCTTTAGAGCATCACCTGTACGTAGTCAGTACGTAAAT 382
Db 524 PheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSerTyrValAsn 543
QY 383 CCTGGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACATTTCTTGTGATCAGTCAG 442
Db 544 ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCysIleSerGln 563
QY 443 CACTGTGACTTCTTTATAGTAAGTATAGTAACACAGAGAGATCCACACTGTGTCCCTT 502
Db 564 HisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeu 583
QY 503 TACAAGCTATCAAGTCCCTGGAAGATGACCCAACTTGCAGAAACAAAGGAATTTTGGCCACC 562
Db 584 TyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPheTrpAlaThr 603
QY 563 ATTTTGGATTGACGAGGTCCTCTTCTGACTATATCTCTCTCCAGAAATTTTCTCTTTGAA 622
Db 604 IleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIlePheSerPheGlu 623
QY 623 AGTACTACTGGATTATACATTTGATGGGATGCTCTACAGGCTCATGATCTACAGCTGGA 682
Db 624 SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly 643
QY 683 ANGAATATCTCTACTGTGCTCTTATATATGTTGGTCTCTCAGGTGAGTGGTGAATAAT 742
Db 644 LysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeuValAsnAsn 663
QY 743 CGGTTAAAGGAGTCAAGTATTTCCGGTTTGAATACCTAGCTCTCTAGTTATGTGTT 802
Db 664 ArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrValVal 683
QY 803 GTAGTGATAGACAAACAGGGATCTGTACCGAGGGCTTAAATTTCAAGCGCCCTTTAA 862
Db 684 ValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLys 703
QY 863 TATAAATGGTCAAATAGAAATTTGACGATCAGGTGGAAGGACTCCCAATATCTAGCTTCT 922
Db 704 TyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyrLeuAlaSer 723
QY 923 CGATATGATTTTCATTGATCTAGATCGTGGGCATCCAGCGCTGCTCTATCGAGGATAC 982
Db 724 ArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyrGlyTyr 743
QY 983 CTCTCCCTGATGGCATTAAATGACAGAGTCAGATATCTTCAGGGTTGCTATTCTCGGGCC 1042
Db 744 LeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAlaGlyAla 763
QY 1043 CCAGTCACTCTGTGGATCTCTATGATACAGATACAGTACAG 1081
Db 764 ProValThrLeuTrpIlePheTyrAspThrGlyTyrThr 776
```

RESULT 2

```
US-11-264-096-1594
; Sequence 1594, Application US/11264096
; Publication No. US2006008479A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF546D1
; CURRENT APPLICATION NUMBER: US/11/264,096
; CURRENT FILING DATE: 2005-11-02
; PRIOR APPLICATION NUMBER: 09/833,245
```



```

; PRIOR APPLICATION NUMBER: US 60/128,705
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: US 09/540,209
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 10444
; SEQ ID NO 6408
; LENGTH: 745
; TYPE: PRT
; ORGANISM: B.fragilis
US-11-079-463-6408

Alignment Scores:
Pred. No.: 6e-29 Length: 745
Score: 372.00 Matches: 117
Percent Similarity: 40.7% Conservative: 77
Best Local Similarity: 24.5% Mismatches: 133
Query Match: 19.2% Indels: 150
DB: 11 Gaps: 17

US-10-825-632-8 (1-1083) x US-11-079-463-6408 (1-745)
QY 23 GGAGTCGCTACCTTTGTTCTCCAGAGAA-----TTGATAGATATCTGGCTATTGG 76
Db 202 GlyIleProAspTrpValTyrGluGluGluPheGlyPheAsnArgAlaLeuGluPhe--- 220
QY 77 TGGTGTCCAAAGCTGAACACACTCCCGGCTGGTGAATAATCTTAGAATTCATATGAA 136
Db 221 -----AsnAlaAspAsnThr-----MetLeuAlaTyrVal 230
QY 137 GMAATGATGATCTGAGGTGGAATTTATTCATGTTACATCCCTATGTTG-----GAA 190
Db 231 ArgPheAspGluSerGluValPro-----SerTyrThrPheProLeuPheAlaGlyGlu 248
QY 191 ACAAGGAGGCGAGAT-----TCATCCGTTATCTTAA 223
Db 249 AlaProArgTyrAspAlaLeuGlnAspTyrProGlyGluTyrThrTyrLysTyrProLys 268
QY 224 ACAGTCAGCAATCTTAAGTC-----ACTTTT----- 253
Db 269 AlaGlyTyrProAsnSerLysValSerValHisThrPheAspIleLysSerLysValThr 288
QY 254 AGAGTCTCAGAAATATGATGATGCTGAGGAGGAGATC----- 292
Db 289 ArgGlnValLysLeuProIleAspAlaAspGlyTyrIleProArgIleArgPheThrGln 308
QY 292 ----- 292
Db 309 AspProAsnLysLeuAlaIleMetThrLeuAsnArgHisGlnAsnArgPheAspMetTyr 328
QY 292 ----- 292
Db 329 PheAlaAspProArgSerThrValCysLysLeuAlaLeuArgAspGluSerProTyrTyr 348
QY 293 -----ATAGTTGAT----- 301
Db 349 IleAsnGluAsnValPheAspAsnIleGlnPheTyrProGluTyrPheSerPheValSer 368
QY 301 ----- 301
Db 369 AspLysSerGlyTyrProHisLeuTyrTrpTyrSerMetAsnGlyLeuLeuLysGln 388
QY 302 -----GAAGTCAGAGGCTGTTA----- 319
Db 389 ValThrSerGlyAsnTyrGluValLysAsnPheIleGlyTyrAsnProAspThrAsnGlu 408
QY 320 ---TATTTTGAAGGCCAACAGACTCCCTTTAGAGCATCCTGTAGCTAGTACAGTTAC 376
Db 409 PheTyrTyrThrSerAsnGluGluSerProMetArgGlnAlaValTyrLysIleAsp--- 427
QY 377 GTAATCTCGAGAGGTGACAGGCTGACTGACCGTGGCTACTCATCTTCTGTGTCATC 436
Db 428 ---ArgLysGlyLysLysMetLysLeuSerAsnGlnProGlyThrAsnSerProIlePhe 446
```

```

QY 437 AGTCAGCACTGTGACTTCTTTATAGTAAGTATAGTAACCAAGAAATCCACACTGTGTG 496
Db 447 SerSerMetLysTyrPheMetAsnLysPheThrSerLeuAspThrProMetLeuIle 466
QY 497 TCCCTTTACAAGCTATCAAGTCTGAGATGACCAACTTGCACAAACAAAGAAATTTGG 556
Db 467 ThrLeu-----AsnAspAsnThrGlyLysValLeuLysThrLeu 479
QY 557 GCCACATTTTGGATTTCAGCAGGTCCTTCTCTGACTATATCTCTCCA-----GAAATT 610
Db 480 ValThrAsnAspLysLeuLysGlnLysLeuAlaGluTyrAlaIleProGlnLysGluPhe 499
QY 611 TTCTCTTTGAAAGTACTACTGATTTTACATTTGATGGATGCTCTACAGCTCATGAT 670
Db 500 PheThrPheLysThrThrGluGlyValAspLeuAsnGlyTyrMetLysProValAsn 519
QY 671 CTACAGCTCGAAGAAATATCTCTACTGCTGCTCTATATATGTTGCTGCTCCTCAGGTGCAG 730
Db 520 PheAspProAlaLysArgTyrProValLeuMetPheGlnTyrSerGlyProGlySerGln 539
QY 731 TTGGTGAATAATCGGTTTAAAGAGTCAAGTATTTCCGGTTGTAATACCTTAGCTCTCTA 790
Db 540 GlnValLeuAspLysTrp---GlyIleSerTrp-----GluThrTyrMetAlaSerLeu 556
QY 791 GGTATGTTGGTTGTAGTATAGACAGACAGGGGATCTCTCACCGAGGCTTAAATTTGAA 850
Db 557 GlyTyrValValAlaCysValAspGlyArgGlyThrGlyArgGlySerGluPheGln 576
QY 851 GCGCGCTTTAAATATAATAATGGTCAAAATAGAAATTCACATGAGTGGAGGACTCCAA 910
Db 577 LysCysThrTyrLeuAsnLeuGlyValLysGluAlaLysAspGlnValGluAlaLys 596
QY 911 TATCTAGCTCTCGATATGATTTCATTTGACTTAGTATGCTGTGGGCATCCACGGCTGTCC 970
Db 597 TyrLeuGlyGly---LeuProTyrValAspLysGlyArgIleGlyIleTrpGlyTyrPse 615
QY 971 TATGAGGATACCTCTCCCTGATGTCATTAATGACAGAGTCAAGATATCTTCAGGTTGCT 1030
Db 616 PheGlyGlyTyrMetThrIleMetSerMetSerGluGlyThrProValPheLysAlaGly 635
QY 1031 ATTGCTGGGCGCCAGTCATCTCTGTGATCTTCTATGATACAGGATACAGC 1081
Db 636 ValAlaValAlaAlaProThrAspTrpLysTyrTyrAspThrValTyrThr 652

RESULT 6
US-11-208-288-4
; Sequence 4, Application US/11208288
; Publication No. US20060051366A1
; GENERAL INFORMATION:
; APPLICANT: CHANG, Chiwen
; TITLE OF INVENTION: USE OF SOLUBLE CD26 AS INHIBITOR OF
; FILE REFERENCE: 39533-0001
; CURRENT FILING DATE: 2005-08-18
; PRIOR APPLICATION NUMBER: US 60/605,013
; PRIOR FILING DATE: 2004-08-26
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 738
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-208-288-4

Alignment Scores:
Pred. No.: 3.03e-23 Length: 738
Score: 315.00 Matches: 120
Percent Similarity: 34.9% Conservative: 59
Best Local Similarity: 23.4% Mismatches: 146
Query Match: 16.3% Indels: 188
DB: 11 Gaps: 22
```

```
US-10-825-632-8 (1-1083) x US-11-208-288-4 (1-738)
QY 2 GAAGAGATCCAGATCAGCTGGAGTCGCTACTTTGTTCTTCCAGAGAA---TTTGAT 58
Db :|||
162 LybGluAspIleIleTyrAsnGlyIleThrAspTyrValTyrGluGluValPheSer 181
QY 59 AGATATTCTCGCTATTGGTGGTGTCCAAAAGCTGAAACCACTCCAGCTGGTGGTAAATT 118
Db :|||
182 AlatySerAlaLeuIleTyrPheProAsnGlyThrPhe----- 194
QY 119 CTAGAAATCTATCAAGAAATGATGATCTGAGTGGAAATTTTCAT----- 169
Db :|||
195 -----LeuAlaTyrAlaGlnPheAsnAspThrGluValProIleIleGluTyrSerPhe 212
QY 170 GTTACATCCCTATGTTGGAACAAGGAGGCGAGTTCATTCCTTATCTCTAAACAGGT 229
Db :|||
213 TyrSerAspGluSerLeuGlnTyrProLysThrValArgValProTyrProLysAlaGly 232
QY 230 ACAGCAAACTCTAAAGTCACTTT----- 253
Db :|||
233 AlaValAsnProThrValLysPheValValAsnThrAspSerLeuSerSerValThr 252
QY 254 AAGATCTCAGAATAATGATGCTCGAAGGAGGATCATAGTTGATGAAGTCAGAAGG 313
Db :|||
253 AsnAlaThrSerIleGlnIleThrAlaProAlaSerMetLeuIleGly----- 268
QY 314 CTGGTATATTTGAAGGCCAACCAAGACTCCCTTTAGAGCATCACCTGTACGTAGTACGT 373
Db :|||
269 -----AspHisTyrLeuCysAspValThr 276
QY 374 TAGCTAAATCTGGAGAGTGACA-----AGGCTGACTGACGCTGGCTAC 418
Db :|||
277 TrpAlaThrGlnGluArgIleSerLeuGlnTrpLeuArgArgIleGlnAsnTyrSerVal 296
QY 419 TCACATCTTTC-----TGCATC-----AGTCAG 442
Db :|||
297 MetAspIleCysAspTyrAspGluSerSerGlyArgTrpAsnCysLeuValAlaArgGln 316
QY 443 CACTGTGACTTC-----TTTATAAGTAAGTATAGTAACTAACAGAGAAATCCACAC 490
Db :|||
317 HisIleGluMetSerThrThrGlyTyrValGlyArgPhe---ArgProSerGluProHis 335
QY 491 TGTGTG-----TCCTTTACAAGCTATCAAGTCCCTGAGAT----- 526
Db :|||
336 PheThrLeuAspGlyAsnSerPheTyrLysIleSerAsnGluGluGlyTyrArgHis 355
QY 527 -----GACCCAACTTGCAAAACAAAGAAATTTGGGCC 559
Db :|||
356 IleCysTyrPheGlnIleAspLysLysAspCysThrPheIleThrLysGlyThrTrpGlu 375
QY 560 ACCATT----- 565
Db :|||
376 ValIleGlyIleGluAlaLeuThrSerAspTyrLeuTyrTyrIleSerAsnGluTyrLys 395
QY 565 ----- 565
Db :|||
396 GlyMetProGlyGlyArgAsnLeuTyrLysIleGlnLeuIleAspTyrThrLysValThr 415
QY 565 ----- 565
Db :|||
416 CysLeuSerCysGluLeuAsnProGluArgCysGlnTyrTyrSerValSerPheSerLys 435
QY 566 -----TTGGATTCCAGCAGCTCTCTCTCTGCTGCTATCTCTCTGCTGCTATCT 598
Db :|||
436 GluAlaLysTyrTyrGlnLeuArgCysSerGlyProGlyLeuProLeuTyrThrLeuHis 455
QY 598 ----- 598
Db :|||
456 SerSerValAsnAspLysGlyLeuArgValLeuGluAspAsnSerAlaLeuAspLysMet 475
QY 599 -----CCTCCAGAAATTTTCTCTTTT-----GAAAGTACTACTGGA 634
Db :|||
476 LeuGlnAsnValGlnMetProSerLysLysLeuAspPheIleIleLeuAsnGluThrLys 495
```

```
635 TTTACATTGTATGGATGCTCTTACAAGCTCTCATGATCTACAGCTGGAAAGAAATATCCT 694
Db :|||
496 Phe---TrpTyrGlnMetIleLeuProProHis---PheAspLysSerLysLysTyrPro 513
QY 695 ACTGTGCTGTATATATGTTGCTCTCAGGTGCAGTTCGTGTGAATAATCGGTTTAAAGGA 754
Db :|||
514 LeuLeuLeuAspValTyrAlaGlyProCysSerGln-----LysAla 527
QY 755 GTCAGATATTTCGCTTGAAT-----ACCTAGCTCTCTAGGTATATGCTGTTGTA 805
Db :|||
528 AspThrValPheArgLeuAsnTrpAlaThrTyrLeuAlaSerThrGluAsnIleVal 547
QY 806 GTG---ATACACAACAGGGATCTCTCACCGAGGCTTAAATTTGAAGGCGCTTTAAA 862
Db :|||
548 AlaserPheAspGlyArgGlySerGlyTyrGlnGlyAspLysIleMetHisAlaIleAsn 567
QY 863 TATAAATGGTCAATAGATAATGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCT 922
Db :|||
568 ArgArgLeuGlyThrPheGluValGluAspGlnIleGluAlaAaArgGlnPhe---Ser 586
QY 923 CGATATGATTTTCATTGACTTAGATCGTGTGGGCATCCAGGCTGTCTATGGAGGATAC 982
Db :|||
587 LysMetGlyPheValAsnAsnLysArgIleAlaIleTrpGlyTyrSerTyrGlyTyr 606
QY 983 CTCTCCCTGATGGCATTATGACAGGTGCAGATATCTTCAGGGTGTCTATGCTGGGGCC 1042
Db :|||
607 ValThrSerMetValLeuGlySerGlySerGlyValPheLysCysGlyIleAlaValAla 626
QY 1043 CAGTCACTCTCTGGATCTCTATGATACAGGATACAG 1081
Db :|||
627 ProValSerArgTrpGluTyrTyrAspSerValTyrThr 639

RESULT 7
US-11-116-939-13
; Sequence 13, Application US/11116939
; Publication No. US2005026595A1
; GENERAL INFORMATION:
; APPLICANT: Stephen Tomlinson
; TITLE OF INVENTION: TISSUE TARGETED COMPLEMENT MODULATORS
; FILE REFERENCE: 19113.0115U2
; CURRENT APPLICATION NUMBER: US/11/116,939
; PRIOR FILING DATE: 2005-04-28
; PRIOR APPLICATION NUMBER: 60/565,907
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 762
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence; note=synthetic
US-11-116-939-13

Alignment Scores:
Pred. No.: 3,046-23 Length: 762
Score: 315.00 Matches: 120
Percent Similarity: 34.9% Conservative: 59
Best Local Similarity: 23.4% Mismatches: 146
Query Match: 16.3% Indels: 188
DB: 11 Gaps: 22

US-10-825-632-8 (1-1083) x US-11-116-939-13 (1-762)
QY 2 GAAGAGATCCAGATCAGCTGGAGTCGCTACTTTGTTCTTCCAGAGAA---TTTGAT 58
Db :|||
186 LysGluAspIleIleTyrAsnGlyIleThrAspTyrValTyrGluGluValPheSer 205
QY 59 AGATATTCTCGCTATTGGTGGTGTCCAAAAGCTGAAACCACTCCAGCTGGTGGTAAATT 118
Db :|||
```

755 GTCAGTATTTCCGCTTGAAT-----ACCTAGCCTCTCTAGTTATGTGTGTA 805
Db 552 AspThrValPheArgLeuAsnTrpAlaThrTyrLeuAlaSerThrGluAsnIleVal 571
Qy 806 GTG---APAGACACAGGGATCTGTCCAGGAGCTTAAATTTGAAGCCCTTTAA 862
Db 572 AlaSerPheAspGlyArgGlySerGlyTyrGlnGlyAspGlyMetHisAlaIleAsn 591
Qy 863 TATAAATGGCTCAATAGAAATGACATCAGGTGGAAGGACTCCCAATATCTAGCTTCT 922
Db 592 ArgArgLeuGlyThrPheGluValGluAspGlnIleGluAlaAlaArgGlnPhe---Ser 610
Qy 923 CCATATGATTTTCATTTAGATTAGATCGTGTGGGCATCCACGGCTGCTCTATGAGGATAC 982
Db 611 LysMetGlyPheValAspAsnLysArgIleAlaIleTrpGlyTrpSerTyrGlyGlyTyr 630
Qy 983 CTCTCCCTGATGGCATTATTCAGAGGTTCAGATATCTTCAGGGTGTCTATTGCTGGGCG 1042
Db 631 ValThrSerMetValLeuGlySerGlySerGlyValPheLysCysGlyIleAlaValAla 650
Qy 1043 CCAGTCACTCTGTGATCTTCTATCATACAGGATACACG 1081
Db 651 ProValSerArgTrpGluTyrTyrAspSerValTyrThr 663

RESULT 8
US-10-522-789-2
; Sequence 2, Application US/10522789
; Publication No. US20050260732A1
; GENERAL INFORMATION:
; APPLICANT: TANABE SEIYAKU CO., LTD.
; TITLE OF INVENTION: Three-dimensional structure of dipeptidyl peptidase IV
; FILE REFERENCE: 03-039-PCT
; CURRENT APPLICATION NUMBER: US/10/522,789
; CURRENT FILING DATE: 2005-01-28
; PRIOR APPLICATION NUMBER: US 60/398,761
; PRIOR FILING DATE: 2002-07-29
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 2
; LENGTH: 766
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-522-789-2

Alignment Scores:
Pred. No.: 3,056-23 Length: 766
Score: 315.00 Matches: 120
Percent Similarity: 34.9% Conservative: 59
Best Local Similarity: 23.4% Mismatches: 146
Query Match: 16.3% Indels: 188
DB: Gaps: 22

US-10-825-632-8 (1-1083) x US-10-522-789-2 (1-766)

Qy 2 GAAGACATCCAGATCAGCTGAGTCGCTACCTTTGTTCTCCAGAGAA---TTTGAT 58
Db 190 LysGluAspIleLeuTyrAsnGlyIleThrAspTrpValTyrGluGluValPheSer 209
Qy 59 AGATATTCTGGCTATTGTTGGTGTCCAAAAGCTGAAACAACTCCAGTGGTGGTAAATTT 118
Db 210 AlaTyrSerAlaLeuTrpTrpSerProAsnGlyThrPhe----- 222
Qy 119 CTTAGAATCTTATATGAGAAATGATGATCTGAGGTGGAATATTTCAT----- 169
Db 223 -----LeuAlaTyrAlaGlnPheAsnAspThrGluValProLeuIleGluTyrSerPhe 240
Qy 170 GTTACATCCCTTATGTTGGAAACAGAGGGCAGATTTCATTCCTGTTATCTTAAACAGGT 229
Db 241 TyrSerAspGluSerLeuGlnTyrProLysThrValArgValProTyrProLysAlaGly 260
Qy 230 ACAGCAATCTTAAAGTCACTTTT----- 253

206 AlaTyrSerAlaLeuTrpTrpSerProAsnGlyThrPhe----- 218
Qy 119 CTTAGAATCTTATATGAGAAATGATGATCTGAGGTGGAATATTTCAT----- 169
Db 219 -----LeuAlaTyrAlaGlnPheAsnAspThrGluValProLeuIleGluTyrSerPhe 236
Qy 170 GTTACATCCCTTATGTTGGAAACAGAGGGCAGATTTCATTCCTGTTATCTTAAACAGGT 229
Db 237 TyrSerAspGluSerLeuGlnTyrProLysThrValArgValProTyrProLysAlaGly 256
Qy 230 ACAGCAATCTTAAAGTCACTTTT----- 253
Db 257 AlaValAsnProThrValLysPhePheValValAsnThrAspSerLeuSerSerValThr 276
Qy 254 AAGATGTCAGAAATAATGATTCATGTCGAAGGAGGATCATGATGTCATGAGTCAGAAGG 313
Db 277 AsnAlaThrSerIleGlnIleThrAlaProAlaSerMetLeuIleGly----- 292
Qy 314 CTGGTATATTTTGAAGGCACCAAGACTCCCTTTAGACATCACCCTGTACGTAGTCAGT 373
Db 293 -----AspHisTyrLeuCysAspValThr 300
Qy 374 TAGCTAAATCTGGAGAGGTGACA-----AGGCTGACTGACCGTGGCTAC 418
Db 301 TrpAlaThrGlnGluArgIleSerLeuGlnTrpLeuArgArgIleGlnAsnTyrSerVal 320
Qy 419 TCACATTTCTTGC-----TGCATC-----AGTCAG 442
Db 321 MetAspIleCysAspTyrAspGluSerSerGlyArgTrpAsnCysLeuValAlaArgGln 340
Qy 443 CACTGTGACTTC-----TTTATAAGTAAATAGTAAACAGAGGAATCCACAC 490
Db 341 HisIleGluMetSerThrThrGlyTrpValGlyArgPhe---ArgProSerGluProHis 359
Qy 491 TGTGTG-----TCCCTTTACAGCTCAAGTCTCTCAAGAT----- 526
Db 360 PheThrLeuAspGlyAsnSerPheTyrLysIleSerAsnGluGluGlyTyrArgHis 379
Qy 527 -----CACCAACTGCAAAACAGAGGAATTTGGGCG 559
Db 380 IleCysTyrPheGlnIleAspLysLysCysPheThrPheIleThrLysGlyThrTrpGlu 399
Qy 560 ACCATT----- 565
Db 400 ValIleGlyIleGluAlaLeuThrSerAspTyrLeuTyrTyrIleSerAsnGluTyrLys 419
Qy 565 ----- 565
Db 420 GlyMetProGlyGlyArgAsnLeuTyrLysIleGlnLeuSerAspTyrThrLysValThr 439
Qy 565 ----- 565
Db 440 CysLeuSerCysGluLeuAsnProGluArgCysGlnTyrTyrSerValSerPheSerLys 459
Qy 566 -----TTGGATTACAGAGCTCT-----CTTCTGACTACTACT- 598
Db 460 GluAlaLysTyrTyrGlnLeuArgCysSerGlyProGlyLeuProLeuTyrThrLeuHis 479
Qy 598 ----- 598
Db 480 SerSerValAsnAspLysGlyLeuArgValLeuGluAspAsnSerAlaLeuAspLysMet 499
Qy 599 -----CTCCAGAAATTTTCTCTTT-----GAAAGTACTACTGGA 634
Db 500 LeuGlnAsnValGlnMetProSerLysLysLeuAspPheIleIleLeuAsnGluThrLys 519
Qy 635 TTTACATGTCATGTCATGCTCTACAAAGCTCTGATCTACAGCTCGGAAAGAAATATCTCT 694
Db 520 Phe---TrpTyrGlnMetIleLeuProHis---PheAspLysSerLysLysTyrPro 537
Qy 695 ACTGTGCTGTCATATATGTCCTCAGGTGCAGTGGTGAATATCGGTTAAAGGA 754
Db 538 LeuLeuLeuAspValTyrAlaGlyProCysSerGln-----LysAla 551

Db 261 AlaValAsnProThrValLysPheValValAsnThrAspSerLeuSerSerValThr 280
QY 254 AAGATGTCAGAAATAATGATTGATGCTGAAGGAGGATCATAGTTGATGAAGTCAGAGG 313
Db 281 AsnAlaThrSerIleGlnIleThrAlaProAlaSerMetLeuIleGly----- 296
QY 314 CTGTATATATTTGAAGGCCACCAAGACTCCCTTTAGAGCATCACTGTAGTAGTCAGT 373
Db 297 -----AspHisTyrLeuCysAspValThr 304
QY 374 TAGCTAAATCTCGAGAGGTGACA-----AGGCTAGCTGACCGTGGCTAC 418
Db 305 TrpAlaThrGlnGluArgIleSerLeuGlnTrpLeuArgArgIleGlnAsnTyrSerVal 324
QY 419 TCACATCTCTGC-----TCCATC-----AGTCAG 442
Db 325 MetAspIleCysAspTyrAspGluSerSerGlyArgTrpAsnCysLeuValAlaAaArgGln 344
QY 443 CACTGTGACTTC-----TTTATAAGTAAATAGTATAACCAAGAAAGAAATCCACAC 490
Db 345 HisIleGluMetSerThrThrGlyTrpValGlyArgPhe---ArgProSerGluProHis 363
QY 491 TGTGTG-----TCCCTTTACAAAGCTCAAGTCTCGAAGAT----- 526
Db 364 PheThrLeuAspGlyAsnSerPheTyrLysIleSerAsnGluGluGlyTyrArgHis 383
QY 527 -----GACCCAACTTGCAAAACAAAGAAATTTTGGGCC 559
Db 384 IleCysTyrPheGlnIleAspLysLysAspCysThrPheIleThrLysGlyThrTrpGlu 403
QY 560 ACCATT-----||| 565
Db 404 ValIleGlyIleGluAlaLeuThrSerAspTyrLeuTyrTyrIleSerAsnGluTyrLys 423
QY 565 ----- 565
Db 424 GlyMetProGlyGlyArgAsnLeuTyrLysIleGlnLeuSerAspTyrThrLysValThr 443
QY 565 ----- 565
Db 444 CysLeuSerCysGluLeuAsnProGluArgCysGlnTyrTyrSerValSerPheSerLys 463
QY 566 -----TTGGATTACAGACAGTCCT---CTTCCCTGACTACT--- 598
Db 464 GluAlaLysTyrTyrGlnLeuArgCysSerGlyProGlyLeuProLeuTyrThrLeuHis 483
QY 598 ----- 598
Db 484 SerSerValAsnAspLysGlyLeuArgValLeuGluAspAsnSerAlaLeuAspLysMet 503
QY 599 -----CCTCCAGAAATTTCTCTTTT-----GAAAGTACTACTGGA 634
Db 504 LeuGlnAsnValGlnMetProSerLysLysLeuAspPheIleLeuAsnGluThrLys 523
QY 635 TTTACATTGTATGGATGCTCTACAGCCTCATGATCTACAGCCTGGAAGAAATATCCT 694
Db 524 Phe---TrpTyrGlnMetIleLeuProProHis---PheAspLysSerLysLysTyrPro 541
QY 695 ACTGTGCTGTTCATATATGTTGCTCTCAGGTGCAAGTGGTGAATAATCGGTTTAAAGGA 754
Db 542 LeuLeuLeuAspValTyrAlaGlyProCysSerGln-----LysAla 555
QY 755 GTCAAGTATTTCCGCTTGAAT-----ACCCTAGGCTCTCTAGGTTATGTTGTTGTA 805
Db 556 AspThrValPheArgLeuAsnTrpAlaThrTyrLeuAlaSerThrGluAsnIleVal 575
QY 806 GTG---ATAGAACACAGGGGATCTGTCCACCGAGGCTTAAATTTGAAGGCGCTTTAAA 862
Db 576 AlaSerPheAspGlyArgGlySerGlyTyrGlnGlyAspLysIleMetHisAlaIleAsn 595
QY 863 TATAAAATGGGTCAAAATAGAAATTGACGATCAGGTGGAGGACTCCCAATATCTAGCTTCT 922
Db 596 ArgArgLeuGlyThrPheGluValGluAspGlnIleGluAlaAaArgGlnPhe---Ser 614

QY 923 CGATATGATTTTCATTGACTTAGATCGTGTGGCATCCACGGCTGCTCTATGGAGTAC 982
Db 615 LysMetGlyPheValAspAsnLysArgIleAlaIleTrpGlyTrpSerTyrGlyTyr 634
QY 983 CTCCTCCCTGATGGCATTAAATGACAGGTCAGATATCTTCAGGTTGCTATTGCTGGGCC 1042
Db 635 ValThrSerMetValLeuGluSerGlySerGlyValPheLysCysGlyIleAlaValAla 654
QY 1043 CCAGTCACTCTGTGGTCTTCTATGATACAGATACAG 1081
Db 655 ProValSerArgTrpGluTyrTyrAspSerValTyrThr 667
RESULT 9
US-10-501-035-234
; Sequence 234, Application US/10501035
; Publication No. US20060046249A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: IDENTIFICATION OF POLYNUCLEOTIDES AND POLYPEPTIDE FOR PREDICTING
; TITLE OF INVENTION: ACTIVITY OF COMPOUNDS THAT INTERACT WITH PROTEIN TYROSINE KINASES
; TITLE OF INVENTION: AND/OR PROTEIN TYROSINE KINASE PATHWAYS
; FILE REFERENCE: D0195 PCT
; CURRENT APPLICATION NUMBER: US/10/501,035
; PRIOR FILING DATE: 2004-07-09
; PRIOR APPLICATION NUMBER: US 60/350,061
; NUMBER OF SEQ ID NOS: 795
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 234
; LENGTH: 766
; ORGANISM: Homo sapiens
US-10-501-035-234
Alignment Scores:
Pred. No.: 3,05e-23 Length: 766
Score: 315.00 Matches: 120
Percent Similarity: 34.9% Conservative: 59
Best Local Similarity: 23.4% Mismatches: 146
Query Match: 16.3% Indels: 188
DB: 9 Gaps: 22
US-10-825-632-8 (1-1083) x US-10-501-035-234 (1-766)
QY 2 GAAGAAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAAGAGAA---TTTGAT 58
Db 190 LysGluAspIleIleTyrAsnGlyIleThrAspTrpValTyrGluGluValPheSer 209
QY 59 AGATATTCTGCTATTGTTGTTGTCCTCAAAAGCTGAAACAACTCCCAAGTGGTGAATAAT 118
Db 210 AlaTyrSerAlaLeuTyrTrpSerProAsnGlyThrPhe----- 222
QY 119 CTAGAATTCTATATGAAGAAATGATGATCTGAGGTGGAATTTATTCAT----- 169
Db 223 -----LeuAlaTyrAlaGlnPheAsnAspThrGluValProLeuIleGluTyrSerPhe 240
QY 170 GTTACATCCCCTATGTTGGAAACAAGGAGGCGAGATTCATTCGGTTATCCTAAAACAGGT 229
Db 241 TyrSerAspGluSerLeuGlnTyrProLysThrValArgValProTyrProLysAlaGly 260
QY 230 ACAGCAATCTTAAGCTCACTTT----- 253
Db 261 AlaValAsnProThrValLysPhePheValValAsnThrAspSerLeuSerSerValThr 280
QY 254 AAGATGTCAGAAATAATGATTGATGCTGAAGGAGGATCATAGTTGATGAAGTCAGAAAG 313
Db 281 AsnAlaThrSerIleGlnIleThrAlaProAlaSerMetLeuIleGly----- 296
QY 314 CTGTATATTTTGAAGGCCACCAAGACTCCCTTTAGAGCATCACCCTTTAGAGCTACGTAGTCT 373
Db 297 -----AspHisTyrLeuCysAspValThr 304

```
QY 374 TAGCTAAATCCTGGAGAGTGACA-----AGGCTGACTGACCGTGGCTAC 418
Db 305 TtpAlaThrGlnGluArgIleSerLeuGlnTrpLeuArgArgIleGlnAsnTrpSerVal 324
QY 419 TCACATTCCTGTC-----TGCATC-----AGTCAG 442
Db 325 MetAspIleCysAspTyrAspGluSerSerGlyArgTrpAsnCysLeuValAlaAraGln 344
QY 443 CACTGTGACTTC-----TTTATAAGTAAAGTATAGTAAACCAAGAAATCCACAC 490
Db 345 HisIleGluMetSerThrThrGlyTrpValGlyArgPhe---ArgProSerGluProHis 363
QY 491 TGTGTG-----TCCCTTTACAAAGTATCAAGTCTGAGAT----- 526
Db 364 PheThrLeuAspGlyAsnSerPheTyrLysIleSerAsnGluGlyTyrArgHis 383
QY 527 -----GACCCAACTTGCAAAACAAAGGAATTTGGGCC 559
Db 384 IleCysTyrPheGlnIleAspLysAspCysThrPheIleThrLysGlyThrTrpGlu 403
QY 560 ACCATT----- 565
Db 404 ValIleGlyIleGluAlaLeuThrSerAspTyrLeuTyrTyrIleSerAsnGluTyrLys 423
QY 565 ----- 565
Db 424 GlyMetProGlyGlyArgAsnLeuTyrLysIleGlnLeuIleAspTyrThrLysValThr 443
QY 565 ----- 565
Db 444 CysLeuSerCysGluLeuAsnProGluArgCysGlnTyrTyrSerValSerPheSerLys 463
QY 566 -----TTGATTCTCAGCAGTCTCT-----CTTCTGACTATACT----- 598
Db 464 GluAlaLysTyrTyrGlnLeuArgCysSerGlyProGlyLeuProLeuTyrThrLeuHis 483
QY 598 ----- 598
Db 484 SerSerValAsnAspLysGlyLeuArgValLeuGluAspAsnSerAlaLeuAspLysMet 503
QY 599 -----CCTCCAGAAATTTTCTCTTT-----GAAAGTACTACTGGA 634
Db 504 LeuGlnAsnValGlnMetProSerLysLysLeuAspPheIleIleLeuAsnGluThrLys 523
QY 635 TTTACATTTGATGGGTCTCTACAGCTCATGATCTACAGCTGGAGAAATATCCT 694
Db 524 Phe---TrpTyrGlnMetIleLeuProHis---PheAspLysSerLysLysTyrPro 541
QY 695 ACTGTGCTGTTCATATATGCTGCTCAGGTGCGAGTTGGTGAATAATCGGTTTAAAGGA 754
Db 542 LeuLeuLeuAspValTyrAlaGlyProCysSerGln-----LysAla 555
QY 755 GTCAAGTATTCGCTTGAAT-----ACCTAGCCTCTCTAGGTTATGTGGTTGTA 805
Db 556 AspThrValPheArgLeuAsnTrpAlaThrTyrLeuAlaSerThrGluAsnIleVal 575
QY 806 GTG---ATAGACAACAGGGATCCTGTCACCGAGGCTTAAATTTGAAGCGCTTTAAA 862
Db 576 AlaSerPheAspGlyArgGlySerGlyTyrGlnGlyAspLysIleMetHisAlaIleAsn 595
QY 863 TATAAATGGGTCAAAATAGAAATGACGATCAGGTGGAGGACTCCAATATCTAGCTTCT 922
Db 596 ArgArgLeuGlyThrPheGluValGluAspGlnIleGluAlaAraGlnPhe---Ser 614
QY 923 CGATATGATTTCAATGACTTAGATCGTGTGGGCATCCAGGCTGCTATGGAGGATAC 982
Db 615 LysMetGlyPheValAspAsnLysArgIleAlaIleTrpGlyTrpSerTyrGlyGlyTyr 634
QY 983 CTCTCCCTGATGCATTAATAGCAGAGGTGAGATATCTTCAGGTTGCTATTCCTGGGCC 1042
Db 635 ValThrSerMetValLeuGlySerGlySerGlyValPheLysCysGlyIleAlaValAla 654
QY 1043 CCAGTCACTGTGGATCTCTCTATGATACAGGATACAG 1081
```

```
Db 655 ProValSerArgTrpGluTyrTyrAspSerValTyrThr 667
RESULT 10
US-11-208-288-2
; Sequence 2, Application US/11208288
; Publication No. US20060051366A1
; GENERAL INFORMATION:
; APPLICANT: CHANG, Chiwen
; TITLE OF INVENTION: USE OF SOLUBLE CD26 AS INHIBITOR OF
; FILE REFERENCE: 39533-0001
; CURRENT APPLICATION NUMBER: US/11/208,288
; CURRENT FILING DATE: 2005-08-18
; PRIOR APPLICATION NUMBER: US 60/605,013
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 766
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-208-288-2
Alignment Scores:
Pred. No.: 3,05e-23 Length: 766
Score: 315.00 Matches: 120
Percent Similarity: 34.9% Conservative: 59
Best Local Similarity: 23.4% Mismatches: 146
Query Match: 16.3% Indels: 188
DB: 11 Gaps: 22
US-10-825-632-8 (1-1083) x US-11-208-288-2 (1-766)
QY 2 GAAGAAGATCCAGATCAGCTGGAGTCGCTACTTTGTTCTCCAGAAGAA---TTTGAT 58
Db 190 LysGluAspIleIleTyrAsnGlyIleThrAspTrpValTyrGluGluValPheSer 209
QY 59 AGATATTCTGGCTATTGGTGTCTCCAAAAGCTGAAACAACTCCCGAGTGGTAAAT 118
Db 210 AlaTyrSerAlaLeuTrpTrpSerProAsnGlyThrPhe----- 222
QY 119 CTAGAATCTATATGAAGAAATATGATCTGAGTGGAAATATTATTCAT----- 169
Db 223 -----LeuAlaTyrAlaGlnPheAsnAspThrGluValProLeuIleLutyrSerPhe 240
QY 170 GTTACATCCCTATGTTGGAAACAGGAGGAGATTCATTCGGTTATCTCTAAACAGGT 229
Db 241 TyrSerAspGluSerLeuGlnTyrProLysThrValArgValProTyrProLysAlaGly 260
QY 230 ACAGCAATCTTAAAGTCACCTTT----- 253
Db 261 AlaValAsnProThrValLysPhePheValValAsnThrAspSerLeuSerSerValThr 280
QY 254 AAGATGTCAGAAATAATGATTGATGCTGAAGAGAGATCATAGTTGATGAAGTCAGAGG 313
Db 281 AsnAlaThrSerIleGlnIleThrAlaProAlaSerMetLeuIleGly----- 296
QY 314 CTGGTATATTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACCTGTACGTAGTCAGT 373
Db 297 -----AspHisTyrLeuCysAspValThr 304
QY 374 TACGTAAATCTCGAGAGGTGACA-----AGGCTGACTGACCGTGGCTAC 418
Db 305 TrpAlaThrGlnGluArgIleSerLeuGlnTrpLeuArgArgIleGlnAsnTrpSerVal 324
QY 419 TCACATTCCTGC-----TGCATC-----AGTCAG 442
Db 325 MetAspIleCysAspTyrAspGluSerSerGlyArgTrpAsnCysLeuValAlaAraGln 344
QY 443 CACTGTGACTTC-----TTTATAAGTAAAGTATAGTAAACCAAGAAATCCACAC 490
Db 345 HisIleGluMetSerThrThrGlyTrpValGlyArgPhe---ArgProSerGluProHis 363
```

```
QY 491 TGTGTG-----TCCCTTTACAAGCTATCAAGTCCTCGAAGAT----- 526
Db 364 PheThrLeuAspGlyAsnSerPheTyrLysIleSerAsnGluGluGlyTyrArgHis 393
QY 527 -----GACCCAACCTGCAAAACAAGGAATTTGGGCC 559
Db 384 IleCysTyrPheGlnIleAspLysLysAspCysThrPheIleThrLysGlyThrTrpGlu 403
QY 560 ACCATT----- 565
Db 404 ValIleGlyLeuAlaLeuThrSerAspTyrLeuTyrTyrIleSerAsnGluTyrLys 423
QY 565 ----- 565
Db 424 GlyMetProGlyGlyArgAsnLeuTyrLysIleGlnLeuIleAspTyrThrLysValThr 443
QY 565 ----- 565
Db 444 CysLeuSerCysGluLeuAsnProGluArgCysGlnTyrTyrSerValSerPheSerLys 463
QY 566 -----TTGGATTACAGAGTCTCT----- 598
Db 464 GluAlaLysTyrTyrGlnLeuArgCysSerGlyProGlyLeuProLeuTyrThrLeuHis 483
QY 598 ----- 598
Db 484 SerSerValAsnAspLysGlyLeuArgValLeuGluAspAsnSerAlaLeuAspLysMet 503
QY 599 -----CCTCCAGAAATTTCTCTTTT-----GAAAGTACTACTCGA 634
Db 504 LeuGlnAsnValGlnMetProSerLysLysLeuAspPheIleIleLeuAsnGluThrLys 523
QY 635 TTTACATTGTATGGGATGCTCTACAGCCTCATGATCTACAGCTCGAAGAAATATCCT 694
Db 524 Phe---TrpTyrGlnMetIleLeuProHis---PheAspLysSerLysLysTyrPro 541
QY 695 ACTGTGCTGTTCAATATGTTGGTCCCTCAGTGCAGTTGCTGATATCGGTTTAAAGA 754
Db 542 LeuLeuLeuAspValTyrAlaGlyProCysSerGln-----LysAla 555
QY 755 GTCAAGTATTTCCGTTGAAT-----ACCTAGCCTCTCAGGTATGCTGTTGTA 805
Db 556 AspThrValPheArgLeuAsnTrpAlaThrTyrIleAlaSerThrGluAsnIleIleVal 575
QY 806 GTG---ATAGACAACAGGGATCCTGTCCAGGAGGCTTAATTTGAAGGCGCCTTTAAA 862
Db 576 AlaSerPheAspGlyArgGlySerGlyTyrGlnGlyAspLysIleMetHisAlaIleAsn 595
QY 863 TATAAATGGCTCAATAGAAATGACATCAGTGAAGGACTCCAATATCTAGCTTCT 922
Db 596 ArgArgLeuGlyThrPheGluValGluAspGlnIleGluAlaAlaArgGlnPhe---Ser 614
QY 923 CGATATGATTTCATTGACTTAGATCGTGTGGGCATCCAGCGTGTCTATGGAGGATAC 982
Db 615 LysMetGlyPheValAspAsnLysArgIleAlaIleTrpSerTyrGlyGlyTyr 634
QY 983 CTCTCCCTGATGGCATTAAATGCAGAGGTGAGATATCTTCAGGGTTGCTATTGCTGGGCC 1042
Db 635 ValThrSerMetValLeuGlySerGlySerGlyValPheLysCysGlyIleAlaValAla 654
QY 1043 CCAGTCACTCTGTGGATCTTCTATGATACAGGATACAGC 1081
Db 655 ProValSerArgTrpGluTyrTyrAspSerValTyrThr 667
```

RESULT 11

```
US-11-208-288-6
; Sequence 6, Application US/11208288
; Publication No. US20060051366A1
; GENERAL INFORMATION:
; APPLICANT: CHANG, Chiwen
; TITLE OF INVENTION: USE OF SOLUBLE CD26 AS INHIBITOR OF
; TITLE OF INVENTION: ANGIOGENESIS AND INFLAMMATION
```

```
; FILE REFERENCE: 39533-0001
; CURRENT APPLICATION NUMBER: US/11/208,288
; PRIOR FILING DATE: 2005-08-18
; PRIOR APPLICATION NUMBER: US 60/605,013
; PRIOR FILING DATE: 2004-08-26
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 760
; TYPE: PRT
; ORGANISM: Mus musculus
US-11-208-288-6
```

Alignment Scores:

Pred. No.:	1,21e-22	Length:	760
Score:	309.00	Matches:	128
Percent Similarity:	37.1%	Conservative:	57
Best Local Similarity:	25.7%	Mismatches:	154
Query Match:	15.9%	Indels:	160
DB:	11	Gaps:	23

US-10-825-632-8 (1-1083) x US-11-208-288-6 (1-760)

```
QY 2 GAAGAGATGCCAGATCAGCTGAGTCGCTACCTTTGTCTCCAAGAGAA---TTTGAT 58
Db 184 GUGLunValIleTyrAsnGlyIleThrAspTrpValTyrGluGluValPheGly 203
QY 59 AGATATCTCGCTATTGGTGGTGTCCAAA-----GCTGAACAACACT 100
Db 204 AlatySerAlaLeuTrpTrpSerProAsnAsnThrPheLeuAlaTyrAlaGlnPheAsn 223
QY 101 CCAGTGGTGGTAAATCTTAGAATCTATATGAAGAAATATGATGATCTGAGGTGAA 160
Db 224 AspThrGlyValProLeuIleGluTyrSerPhe---TyrSerAspGluSer----- 239
QY 161 ATTATTCATGTTACATCCCTATGTTGAAACAAGGAGGCGAGATTCATTCGCTTATCCT 220
Db 240 -----LeuGlnTyrProLysThrValTrpIleProTyrPro 251
QY 221 AAAACAGGTACAGCAAACTCTAAAGTCACCTTT----- 253
Db 252 LysAlaGlyAlaValAsnProThrValPhePheIleValaAsnIleAspSerLeuSer 271
QY 253 ----- 253
Db 272 SerSerSerSerAlaAlaProIleGlnIleProAlaProAlaSerValAlaArgGlyAsp 291
QY 254 ---AAGATGTCAGAAATAATGATTGCTGCTGAGGAGGATCATAGTTGATGAGTCAGA 310
Db 292 HisTyrLeuCysAspValValTrpAlaThrGluGluArgIleSerLeuGlnTrpLeuArg 311
QY 311 AGGCTGGTATATTTTGAAGGC-----ACCAAA 337
Db 312 ArgIleGlnAsnTyrSerValMetAlaIleCysAspTyrAspLysIleAsnLeuThrTrp 331
QY 338 GACTCCCTCTTAGAG---CATCACCTGTACGTA-----GTCAGTTACGTA 379
Db 332 AsnCysProSerGluGlnGlnHisValGluMetSerThrThrGlyTrpValGlyArgPhe 351
QY 380 AATCTGGAGAGGTGCAAGCTGACTGAC----- 409
Db 352 ArgProAlaGluProHisPheThrSerAspGlySerSerPheTyrIleIleSerAsp 371
QY 410 ---CGTGGCTACTCATTCTTCTGTCATC-----AGTCAGCAGCTGTGACTTC--- 454
Db 372 LysAspGlyTyrIleLysHisIleCysHisPheProLysAspLysLysAspCysThrPheIle 391
QY 455 -----TTTAAAGTAAG 466
Db 392 ThrLysGlyAlaTrpGluValIleSerIleGluAlaLeuThrSerAspTyrLeuTyrTrp 411
QY 467 TATAGTAACCAAG-----AAGAAATCCACACTGTGTGCTCCCTTTACAAG----- 508
```


Db 412 IleSerAsnGlnTyrLysGluMetProGlyGlyArgAsnLeuTyrLysIleGlnLeuThr 431
Qy 509 -----CTATCAAGTCTGAAAGATGAGTACCCCACTTGGAAA----- 541
Db 432 AspHisThrAsnValLysCysLeuSerCysAspLeuAsnProGluArgCysGlnTyrTyr 451
Qy 542 -----ACAAAGGAATTTTGGGCCACCATTTTGGATTTCAGCAGGTCCT---CTT 586
Db 452 AlaValSerPheSerLysGluAlaLysTyrTyrGlnLeuGlyCysTrpGlyProGlyLeu 471
Qy 587 CTTGACTATACT----- 598
Db 472 ProLeuTyrThrLeuHisArgSerThrAspHisLysGluLeuArgValLeuGluAspAsn 491
Qy 599 -----CTCCAGAAATTTTCTCTTTT--- 619
Db 492 SerAlaLeuAspArgMetLeuGlnAspValGlnMetProSerLysLysLysLeuAspPheIle 511
Qy 620 ---GAAAGTACTACTGGATTTTACATTTGATGGATGCTCTACAAAGCCTCATGATCTACAG 676
Db 512 ValLeuAsnGlnThrArgPhe---TrpTyrGlnMetIleLeuProProHis---PheAsp 529
Qy 677 CTGGGAAGAAATATCTACTGCTGTCTCATATATGTTGCTCTCAGGTCGATTCGGTG 736
Db 530 LysSerLysLysTyrProLeuLeuLeuAspValTyrAlaGlyProCysSerGln----- 547
Qy 737 AATAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAAT-----ACCTAGCCTCT 787
Db 548 -----LysAlaAspAlaSerPheArgLeuAsnTrpAlaThrTyrLeuAlaSer 563
Qy 788 CTAGGTATGTTGTTGATG---ATAGACAACAGGGATCCTGTCACCGAGGCTTAAA 844
Db 564 ThrGluAsnIleIleValAlaSerPheAspGlyArgGlySerGlyTyrGlnGlyAspLys 583
Qy 845 TTTGAAGGCGCTTTAAATATAATAATGGTCAAAATAGAAATTCAGATCAGGTGGAAGGA 904
Db 584 IleMetHisAlaIleAsnArgGlyGlyThrLeuGluValGluAspGlnIleGluAla 603
Qy 905 CTCCAATATCTAGTCTTCGATATGATTTCAATTTAGATTCGTGTGGGATCCACGGC 964
Db 604 AlaArgGlnPheVal---LysMetGlyPheValAspSerLysArgValAlaIleTrpGly 622
Qy 965 TGCTCTATGCGGATACCTCTCCCTGATGGCATTAAATGCAGAGTCAGATATCTCAGG 1024
Db 623 TrpSerTyrGlyGlyTyrValThrSerMetValLeuGlySerGlySerGlyValPheLys 642
Qy 1025 GTTGTCTATTGCTGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACAGC 1081
Db 643 CysGlyIleAlaValAlaProValSerArgTrpGluTyrTyrAspSerValTyrThr 661

RESULT 12

US-11-186-284-55
; Sequence 55, Application US/11186284
; Publication No. US20050266493A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Berger, Allison
; APPLICANT: Guillemette, Tracy L.
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John B.
; APPLICANT: Thibodeau, Stephen N.
; APPLICANT: BURGART, Lawrence J.
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; FILE REFERENCE: MEMO1-029P22RM
; CURRENT APPLICATION NUMBER: US/11/186,284
; CURRENT FILING DATE: 2005-07-21
; PRIOR APPLICATION NUMBER: US/10/301,822
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 60/339,971
; PRIOR FILING DATE: 2001-12-10

; PRIOR APPLICATION NUMBER: US 60/361,978
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/381,988
; PRIOR FILING DATE: 2002-05-20
; NUMBER OF SEQ ID NOS: 228
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 55
; LENGTH: 760
; TYPE: PRN
; ORGANISM: Homo Sapiens
; US-11-186-284-55

Alignment Scores:
Pred. No.: 1,94e-17 Length: 760
Score: 257.00 Matches: 110
Percent Similarity: 36.8% Conservative: 68
Best Local Similarity: 22.7% Mismatches: 158
Query Match: 13.3% Indels: 148
DB: 11 Gaps: 20

US-10-825-632-8 (1-1083) x US-11-186-284-55 (1-760)

Qy 23 GGAGTCGCTACCTTTGTTCTTCCAAAGAAATTT-----GATAGATATTCTGGCTATTGG 76
Db 195 GlyIleProAspTrpValTyrGluGluMetLeuProThrLysTyrAla---LeuTrp 213
Qy 77 TGGTGTCCAAAGCTGAAACAACACTCCCGAGTGGTGTAAATTTCTTAGAATTTCTATATGAA 136
Db 214 TrpSerPro-----AsnGlyLysPheLeu-----AlaTyrAla 224
Qy 137 GAAATGATGATGATCGAGTGGAAATTTATCATGTTTACATCCCTATGTTGGAACA--- 193
Db 225 GluPheAsnAspLysAspIleProValIleAlaTyrSerTyrTyrGlyAspGluGlnTyr 244
Qy 194 AGGAGGCGAGATTCATTCCGTTATCTCTAAACAGGTACAGCAATCCTAAAGTCACCTTT 253
Db 245 ProArgThrIleAsnIleProTyrProLysAlaGlyAlaLysAsnProValValArgIle 264
Qy 254 AAGATG-----TCAGAAATAATGATGATGCT 280
Db 265 PheIleIleAspThrThrTyrProAlaTyrValGlyProGlnGluValProValProAla 284
Qy 281 -----GAAGGAGGATC 292
Db 285 MetIleAlaSerSerAspTyrTyrPheSerTrpLeuThrTrpValThrAspGluArgVal 304
Qy 293 ATAGTTGATGAAGTCAGAGGCTGGTATATTTTGA--- 328
Db 305 CysLeuGlnTrpLeuLysArgValGlnAsnValSerValLeuSerIleCysAspPheArg 324
Qy 329 -----GGCACCACAAAGACTCCCT---TTAGACATCACCCTGTAC--- 364
Db 325 GluAspTrpGlnThrTrpAspCysProLysThrGlnGluHisIleGluSerArgThr 344
Qy 365 -----GTAGTCAGTTACGTAAATCCTGGAGAG 391
Db 345 GlyTrpAlaGlyGlyPhePheValSerArgProValPheSerTyrAspAlaIleSerTyr 364
Qy 392 GTGCAAGGCTGACTGACCGT---GGCTACTCATCTTCTGTGTCATC----- 436
Db 365 TyrIysIlePheSerAspLysAspGlyTyrLysHisIleHisTyrIleLysAspThrVal 384
Qy 437 -----AGT 439
Db 395 GluAsnAlaIleGlnIleThrSerGlyLysTyrGluAlaIleAsnIlePheArgValThr 404
Qy 440 CAGCACTGTGACTTCTTTTATAAGT-----AAGTATAGTACCAAGAGATCCA 487
Db 405 GlnAspSerLeuPheTyrSerSerAsnGluPheGluGluTyrProGlyArgArgAsnIle 424
Qy 488 CACTGTGTGCTCCCTTTTACAAGCTATCAGTCTCGAAGATGACCCCACTTGCACAAACAG 547
Db 425 TyrArgIleSerIleGlySerTyrProProSerLysLysCysValThrCysHisLeuArg 444

```
QY 548 GAA-----TTTGGGCCACCATTTTGGATTACGCA----- 577
Db 445 Ly8GluArgCysGlnTyrThrAlaSerPheSerAspTyrAlaLysTyrThrAlaLeu 464
QY 578 -----GGTCCT---CTTCTGACTATCT----- 598
Db 465 ValCysTyrGlyProGlyLeuPheLeuHisAspGlyValThrAspGlnGlu 484
QY 599 ----- 601
Db 485 IleLysIleLeuGluGluAenLysGluLeuGluAenAlaLeuLysAenIleGlnLeuPro 504
QY 602 CCAAAATTTCTCTTTTGAAGTACTACTGGATTACATTTGATGGGATGCTCTCAAG 661
Db 505 LysGluGluIleLysLysLeuGluValAspGluLeuThrLeuThrLeuThrLysMetIleLeu 524
QY 662 CCTCATGATCTACAGCTGGAAAGAAATATCTACTGTGCTGTTCATATATGTGTCT 721
Db 525 ProProGlnPheAspArgSerLysLysTyrProLeuLeuIleGlnValTyrGlyGlyPro 544
QY 722 CAGGTGCAGTTGGTGAATTCGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTTA 781
Db 545 CysSerGlnSerValArgSer-----ValPheAlaValAenTrpIle 558
QY 782 GCCTCTCTA-----GGTTATGTGTTGTGATGATACACACAGGGATCCTGT 829
Db 559 SerTyrLeuAlaSerLysGluGlyMetValIleAlaLeuValAspGlyArgGlyThrAla 578
QY 830 CACCCAGGGCTTAATTTGAAGGCGCTTTAAATATATAATATGCTCAAAATAGAAATTGAC 889
Db 579 PheGlnGlyAspLysLeuLeuTyrAlaValTyrArgLysLeuGlyValTyrGluValGlu 598
QY 890 GATCAGGTGGAGGAGTCCAAATATCTAGCTTCTCGATATGATTTCATTGACTTAGATCGT 949
Db 599 AspGlnIleThrAlaValArgLysPheIle---GluMetGlyPheIleAspGluLysArg 617
QY 950 GTGGCATCCACGCTGCTCTATGAGGATACCTCTCTCTGATGATGATGATGATGATGATG 1009
Db 618 IleAlaIleThrAlaValArgLysPheIle---GluMetGlyPheIleAspGluLysArg 617
QY 1010 TCAGATATCTTCAGGTTGTATTGCTGGGCGCCAGTCACTGTGGATCTCTCTATCAT 1069
Db 638 ThrGlyLeuPheLysCysGlyIleAlaValAlaProValSerSerTyrGlyTyrAla 657
QY 1070 ACAGGATACACG 1081
Db 658 SerValTyrThr 661

RESULT 13
; Sequence 7504, Application US/11079463
; Publication No. US20060073161A1
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTERIOIDES FRA
; FILE REFERENCE: FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/11/079,463
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/128,705
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: US 09/540,209
; NUMBER OF SEQ ID NOS: 10444
; SEQ ID NO 7504
; LENGTH: 624
; TYPE: PRT
; ORGANISM: B. fragilis
US-11-079-463-7504

Alignment Scores: 1.19e-11 Length: 624
Pred. No.: 624
```

```
Score: 199.00 Matches: 93
Percent Similarity: 37.9% Conservative: 50
Best Local Similarity: 24.7% Mismatches: 140
Query Match: 10.3% Indels: 94
DB: 11 Gaps: 18
US-10-825-632-8 (1-1083) x US-11-079-463-7504 (1-624)
QY 56 GATAGATATTCTGGCTATTGGTGTCTCCAAAAGCTGAAACAACTCCAGTGGTGGTAAA 115
Db 189 AspArgPheLeuGly-----GlyAlaSerTyrSerProAspGlyLysGln 203
QY 116 ATCTCTAGAATTTCTATATGAAGAA----- 139
Db 204 LeuLeuLeuThrAlaSerProGluAlaPheAspGlyIleGlyLysAsnCysGlyAsnHis 223
QY 140 -----AATGAT---CAATCTGAGTGGGAAATATTATTCATGTTACATCCCTTATGTTG 187
Db 224 ProIleAlaAsnAspPheAspSerGlnAlaPheIleMetAsp-----Leu 238
QY 188 GAAACAAGAGGAGGAGATTCATTCGTTATCTCTAAACAGGTACAGCAATCTCTAAAGTC 247
Db 239 AlaThrArgLysIleAsp-----ProIleThrLysGluPheAsnProSerVal 254
QY 248 ACTTTTAAGATGTCAGAAATAATGATGCTGTGAGGAAGGATCATAGTT----- 298
Db 255 AsnPheLeuGlnTrpAsn-----LysGlyAspGlyCysIleTyrPheSerThrAsn 271
QY 299 GATGAGTACAGAGGCTGGTATATTTTGAAGGACCAAGAACTCCCTTTAGAGCAT--- 355
Db 272 AspGluAspCysArgAsnIleTyrArgTyrSerProLysAspArgLysPheGluLysLeu 291
QY 356 -----CACCTGTAAGTACGTACGTACGTACGTACGTACGTACGTACGTACGTACGT 394
Db 292 AsnLeuGluThrAspValThrSerAlaPheAlaMetSerGluAsnAsnPro----- 308
QY 395 ACAAGGCTGACTGACCGGTGCTACTCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 454
Db 309 -----SerLeuAlaAlaTyrIleGlyGlnGlyCys----- 318
QY 455 TTTATAAGTAAGTATAGTAAACAGAGAAATTTGGGCCACCACTTTTGGATTCTCA 514
Db 319 -----TyrAsnAlaGlyValAlaTyrValTyrAspLeuLys 330
QY 515 AGTCCTGAAGATGACCAACTTGCANAAACAAAGGAATTTTGGGCCACCACTTTTGGATTCTCA 574
Db 331 Lys-----LysThrSerArgLeu-----IleAlaAspPro 340
QY 575 GCAGGTCTCT-----CTTCCTGACTATATCTCTCCAGAAATTTTCTCT 616
Db 341 MetLysProThrLeuGluLysIleGluLeuGlyGluMetLysPro-----TrpAsn 357
QY 617 TTTGAAAGTACTACTGGATTATACATTTGATGGATGCTCTACAGGCTCATGATCTACAG 676
Db 358 PheThrAlaSerAspGlyThrGluIleLysGlyMetMetCysLeuProProSerPheAsp 377
QY 677 CCTGAAAGAAATATCTCTGCTGCTGCTCATATATGTTGTT-----CCTCAGGTGAG 730
Db 378 ProAsnLysLysTyrProLeuIleValTyrTyrGlyGlyThrThrProThrGluArg 397
QY 731 TTGTTGAATAATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCTGCTCTCTA 790
Db 398 GlyIleSerAsnProTyrCysAla-----GlnLeuPheAlaSerArg 411
QY 791 GGTATGTGTTGTTGATGATAGACACAGGGATCTGTCCAGGAGGCTTTAAATTTGAA 850
Db 412 AspTyrValValTyrValIleGlnProSerGlyThrIleGlyPheGlyGlnGluPheSer 431
QY 851 GCGCGCTTTAAATATATAAATGGTCAATAGAAATTTGACATACAGTACAGGTGGAAGGCTCAA 910
Db 432 AlaArgHisValAsnAlaTrpGlyLysArgThrAlaAspAspIleIleGlyGlyThrLys 451
QY 911 TATCTAGCTTCTCGATATGATTTTCATTGACTGCTGTGGGCATCCACGGCTGTGCTCC 970
```

```
Db 452 GlnPheCysLysGluHisProPheValAsnAspLysLysIleGlyCysLeuGlyAlaSer 471
Qy 971 TATGAGGATACCTCCCTGATGCATTAATGACAGAGGTACAGATATCTTCAGGGTTGCT 1030
Db 472 TyrGlyGlyPheMetThr---GlnTyrLeuGlnThrGlnThrAspIlePheAlaAlaA 490
Qy 1031 ATTGCTGGGGCCCCAGTCACT---CTGGGATCTTCTATGATACAGGATAC 1078
Db 491 ValSerHisAlaGlyIleSerAsnValThrSerTyrTrpGlyGluGlyTyr 507

RESULT 14
US-11-179-977-1
; Sequence 1, Application US/11179977
; Publication No. US20050249789A1
; GENERAL INFORMATION:
; APPLICANT: Genencor International, Inc.
; TITLE OF INVENTION: Alpha/Beta Hydrolase-Fold Enzymes
; FILE REFERENCE: GC511-PCT
; CURRENT APPLICATION NUMBER: US/11/179,977
; CURRENT FILING DATE: 2005-07-12
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 657
; TYPE: PRT
; ORGANISM: Bacillus
US-11-179-977-1

Alignment Scores:
Pred. No.: 2,39e-11 Length: 657
Score: 196.00 Matches: 69
Percent Similarity: 40.4% Conservative: 43
Best Local Similarity: 24.9% Mismatches: 107
Query Match: 10.1% Indels: 58
DB: 11 Gaps: 9

US-10-825-632-8 (1-1083) x US-11-179-977-1 (1-657)
Qy 332 ACCAAGACTCCCTTTAGAGTCACCTGATGACGTAGTACGTAAATCTCGGAGAG 391
Db 311 ThrLysAspSer-----GlnGlyPheTyrValIleGly----- 321
Qy 392 GTGACAAGCTCACTGACCGTGC-----TACTCACAATCTTCTGCTGCATCAGTCAGTGT 415
Db 322 -----ThrAspGlnGlySerThrGlyIleTyrIleSerIleGluGlyLeu 337
Qy 416 -----TACTCACAATCTTCTGCTGCATCAGTCAGTGT 448
Db 338 ValTyrProIleArgLeuGluLysGluTyrIleAsnSerPheSerLeuSerProAspGlu 357
Qy 449 GACTCTCTTTAAGTAAATAGTAAACAGAGAAATCCACACTGTGTGTCCTTTTACAAG 508
Db 358 GlnHisPheIleAlaSerValThrLysProAspArgProSerGluLeuTyrSerIlePro 377
Qy 509 CTATCAAGTCTCCAGATGACCCCACTTGCACCAAGAAAGATTTTGGGCCACCAITTG 568
Db 378 LeuGlyGlnGluLysGlnLeuThrGlyAlaAsnAspLysPhe----- 392
Qy 569 GATTTCAGCAGGCTCTCTCCTCACTATACT-----CCTCCAGAAATTTTCTCTTTTGA 622
Db 393 -----ValArgGluHisThrIleSerIleProGluGluIleGlnTyrAla 407
Qy 623 AGTACTACTGGATTTACATTTGATGCGATGCTTACAGCCTCATGATCTACAGCCTGGA 682
Db 408 ThrGluAspGlyValMetValAsnGlyTrpLeuMetArgProAlaGlnMetGluGlyGlu 427
Qy 683 AAGAAATATCTACTGTGCTGTATATATGCTGCTCCTCAGGTGCGAGTTGGTGAATAAT 742
Db 428 ThrThrTyrProLeuIleLeuAsnIleHisGlyGlyProHisMetMetTyr----- 444
Qy 743 CGGTTTAAAGGAGTCAAGTATTTC---CGCTTGAATACCTAGCCTCTCTAGTTATGTG 799
```

```
Db 445 -----GlyHisThrTyrPheHisGluPheGlnValLeuAlaAlaLysGlyTyrAla 461
Qy 800 GTTGTAGTGATAGACAACAGGGATCTCTGTCCACGAGGCTTAAATTTGAAGCGCCTTT 859
Db 462 ValValTyrIleAsnProArgLysSerHisGlyTyrGlnGluPheValAsnAlaVal 481
Qy 860 AAATATAAATGGTCCAAATAGAAATTGACGATCAGGTGGAAGGACTCCAATATCTAGCT 919
Db 482 ArgGlyAspTyrGlyGlyLysAspTyrAspAspValMetGlnAlaValAspGluAlaIle 501
Qy 920 TCTCGATATGATTTCATTGACTTAGTCTGTGGGCATCCACGGCTGCTCTATGGAGGA 979
Db 502 LysArgAspProHisIleAspProLysArgLeuGlyValThrGlyGlySerTyrGlyGly 521
Qy 980 TACCTCTCCCTGATGCATTAAATGACAGAGTCCAGATATCTTCAGGGTTGCTATTGCTGG 1039
Db 522 PheMetThrAsnTrp---IleValGlyGlnThrAsnArgPheLysAlaAlaValThrGln 540
Qy 1040 GCCCCAGTCACTCTGTGGATCTTCTAT-----GATACAGGATAC 1078
Db 541 ArgSerIleSerAsnTrpIleSerPheHisGlyValSerAspIleGlyTyr 557

RESULT 15
US-11-264-096-1591
; Sequence 1591, Application US/11264096
; Publication No. US20060084794A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF546D1
; CURRENT APPLICATION NUMBER: US/11/264,096
; CURRENT FILING DATE: 2005-11-02
; PRIOR APPLICATION NUMBER: 09/833,245
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 2267
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1591
; LENGTH: 349
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (183)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (191)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (192)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (334)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (344)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (345)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (348)
```

! OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-11-264-096-1591

Alignment Scores:
Pred. No.: 3.35e-11 Length: 349
Score: 194.00 Matches: 37
Percent Similarity: 69.4% Conservative: 13
Best Local Similarity: 51.4% Mismatches: 22
Query Match: 10.0% Indels: 0
DB: 11 Gaps: 0

US-10-825-632-8 (1-1083) x US-11-264-096-1591 (1-349)

QY	263	GAAATATGATGATGCTGAAGAGGATCATAGTTGATGAGTCAGAGGCTGGTATAT	322
DB	129	GluValLeuAlaArgHisGlySerLysIleTrpValAsnGluGluThrLysLeuValTyr	148
QY	323	TTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGACGTACGTACGTAAT	382
DB	149	PheGlnGlyThrLysAspThrProLeuGluHisLeuTyrValValSerTyrGluAla	168
QY	383	CCTGGAGAGGTGACAGGCTGACTGACCGTGGCTACTCATCTTCTGTCATCAGT	442
DB	169	AlaGlyGluIleValArgLeuThrThrProGlyPheSerHis**CysSerMetSerGln	188
QY	443	CACTGTGACTTCTTATAGTATAGTAAACAG	478
DB	189	AsnPhe*****PheValSerHisIleThrAlaGln	200

Search completed: May 2, 2006, 03:10:58
Job time : 38.1003 secs

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☒ **BLACK BORDERS**
- ☒ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKewed/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☐ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.

THIS PAGE BLANK (USPTO)